

*Sub
R1*

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: MEDICAL RESEARCH COUNCIL
- (B) STREET: 20 PARK CRESCENT
- (C) CITY: LONDON
- (E) COUNTRY: UK
- (F) POSTAL CODE (ZIP): W1N 4AL

(ii) TITLE OF INVENTION: GENE

(iii) NUMBER OF SEQUENCES: 16

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 924 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 5..604

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TGTC ATG TTG CGG GCT TTG AAC CGC CTG GCC GCG CGG CCC GGG GGC CAG
49

Met Leu Arg Ala Leu Asn Arg Leu Ala Ala Arg Pro Gly Gly Gln
1 5 10 15

CCC CCA ACC CTG CTC CTT CTG CCC GTG CGC GGC CGC AAG ACC CGC CAC
97

Pro Pro Thr Leu Leu Leu Pro Val Arg Gly Arg Lys Thr Arg His
20 25 30

GAT CCG CCT GCC AAG TCC AAG GTC GGG CGC GTG AAA ATG CCT CCT GCA
145

Asp Pro Pro Ala Lys Ser Lys Val Gly Arg Val Lys Met Pro Pro Ala
35 40 45

GTG GAC CCT GCG GAA TTG TTC GTG TTG ACC GAG CGC TAC CGA CAG TAC
193

Val Asp Pro Ala Glu Leu Phe Val Leu Thr Glu Arg Tyr Arg Gln Tyr
50 55 60

CGG GAG ACG GTG CGC GCT CTC AGG CGA GAG TTC ACA TTG GAG GTG CGA
241

Arg Glu Thr Val Arg Ala Leu Arg Arg Glu Phe Thr Leu Glu Val Arg
65 70 75

GGG AAA TTG CAC GAG GCC CGA GCC GGG GTT CTG GCT GAG CGC AAG GCG
289

Gly Lys Leu His Glu Ala Arg Ala Gly Val Leu Ala Glu Arg Lys Ala
80 85 90 95

CAA GAG GCC ATC AGA GAG CAC CAG GAG CTG ATG GCC TGG AAC CGG GAG
337

Gln Glu Ala Ile Arg Glu His Gln Glu Leu Met Ala Trp Asn Arg Glu
100 105 110

GAG AAC CGG AGA CTG CAG GAA CTA CGG ATA GCT AGG TTG CAG CTC GAA
385

Glu Asn Arg Arg Leu Gln Glu Leu Arg Ile Ala Arg Leu Gln Leu Glu
115 120 125

GCA CAG GCC CAG GAG CTG CGG CAG GCT GAG GTC CAG GCC CAG AGG GCC

433

Ala Gln Ala Gln Glu Leu Arg Gln Ala Glu Val Gln Ala Gln Arg Ala
130 135 140

CAG GAG GAG CAG GCT TGG GTG CAA CTG AAA GAA CAA GAA GTT CTC AAA

481

Gln Glu Glu Gln Ala Trp Val Gln Leu Lys Glu Gln Glu Val Leu Lys
145 150 155

CTG CAG GAG GAG GCC AAA AAC TTC ATC ACT CGG GAG AAC CTG GAG GCA

529

Leu Gln Glu Glu Ala Lys Asn Phe Ile Thr Arg Glu Asn Leu Glu Ala
160 165 170 175

CGG ATA GAA GAG GCC TTG GAC TCT CCG AAG AGT TAT AAC TGG GCG GTC

577

Arg Ile Glu Glu Ala Leu Asp Ser Pro Lys Ser Tyr Asn Trp Ala Val
180 185 190

ACC AAA GAA GGG CAG GTG GTC AGG AAC TGAGAACAGA GGCCTCTCAG

624

Thr Lys Glu Gly Gln Val Val Arg Asn
195 200

GCCCCAAATAA GGACAGTGCT TGCCTAGGGA CTGGATATTG GGGTAGAAAT
TGGTGCACTCC 684

CAGGAGGGTG GCACAGCCTT GTCCAGAGCA GCCCCCATTG ATTCTAGATT
TGGCACCAAGG 744

TATAGTACCT GTTCTGACAC CACATACAAA CTCCGGACAG CATTAAACTC
TGGGAAGTTC 804

CTATCACACA GAAGATCAGA CTGGACTGTC CCCTCTAGAA GCCAAGAGCT
GTCTCCTGAG 864

TTTCTTGGAA TAGTGTGAGC CCAATGTTTC CTGCTTTAT AAATAAACTA
TTGGAAAGCA 924

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 200 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Leu Arg Ala Leu Asn Arg Leu Ala Ala Arg Pro Gly Gly Gln Pro
1 5 10 15

Pro Thr Leu Leu Leu Pro Val Arg Gly Arg Lys Thr Arg His Asp
20 25 30

Pro Pro Ala Lys Ser Lys Val Gly Arg Val Lys Met Pro Pro Ala Val
35 40 45

Asp Pro Ala Glu Leu Phe Val Leu Thr Glu Arg Tyr Arg Gln Tyr Arg
50 55 60

Glu Thr Val Arg Ala Leu Arg Arg Glu Phe Thr Leu Glu Val Arg Gly
65 70 75 80

Lys Leu His Glu Ala Arg Ala Gly Val Leu Ala Glu Arg Lys Ala Gln
85 90 95

Glu Ala Ile Arg Glu His Gln Glu Leu Met Ala Trp Asn Arg Glu Glu
100 105 110

Asn Arg Arg Leu Gln Glu Leu Arg Ile Ala Arg Leu Gln Leu Glu Ala
115 120 125

Gln Ala Gln Glu Leu Arg Gln Ala Glu Val Gln Ala Gln Arg Ala Gln
130 135 140

Glu Glu Gln Ala Trp Val Gln Leu Lys Glu Gln Glu Val Leu Lys Leu
145 150 155 160

Gln Glu Glu Ala Lys Asn Phe Ile Thr Arg Glu Asn Leu Glu Ala Arg
165 170 175

Ile Glu Glu Ala Leu Asp Ser Pro Lys Ser Tyr Asn Trp Ala Val Thr
180 185 190

Lys Glu Gly Gln Val Val Arg Asn
195 200

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 998 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:1..615

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATG CTA CGC GCG CTG AGC CGC CTG GGC GCG GGG ACC CCG TGC AGG CCC
48
Met Leu Arg Ala Leu Ser Arg Leu Gly Ala Gly Thr Pro Cys Arg Pro
205 210 215

CGG GCC CCT CTG GTG CTG CCA GCG CGC GGC CGC AAG ACC CGC CAC GAC
96
Arg Ala Pro Leu Val Leu Pro Ala Arg Gly Arg Lys Thr Arg His Asp
220 225 230

CCG CTG GCC AAA TCC AAG ATC GAG CGA GTG AAC ATG CCG CCC GCG GTG
144
Pro Leu Ala Lys Ser Lys Ile Glu Arg Val Asn Met Pro Pro Ala Val
235 240 245

GAC CCT GCG GAG TTC TTC GTG CTG ATG GAG CGT TAC CAG CAC TAC CGC
192

Asp Pro Ala Glu Phe Phe Val Leu Met Glu Arg Tyr Gln His Tyr Arg
 250 255 260

CAG ACC GTG CGC GCC CTC AGG ATG GAG TTC GTG TCC GAG GTG CAG AGG
 240
 Gln Thr Val Arg Ala Leu Arg Met Glu Phe Val Ser Glu Val Gln Arg
 265 270 275 280

AAG GTG CAC GAG GCC CGA GCC GGG GTT CTG GCG GAG CGC AAG GCC CTG
 288
 Lys Val His Glu Ala Arg Ala Gly Val Leu Ala Glu Arg Lys Ala Leu
 285 290 295

AAG GAC GCC GCC GAG CAC CGC GAG CTG ATG GCC TGG AAC CAG GCG GAG
 336
 Lys Asp Ala Ala Glu His Arg Glu Leu Met Ala Trp Asn Gln Ala Glu
 300 305 310

AAC CGG CGG CTG CAC GAG CTG CGG ATA GCG AGG CTG CGG CAG GAG GAG
 384
 Asn Arg Arg Leu His Glu Leu Arg Ile Ala Arg Leu Arg Gln Glu Glu
 315 320 325

CGG GAG CAG GAG CAG CGG CAG GCG TTG GAG CAG GCC CGC AAG GCC GAA
 432
 Arg Glu Gln Glu Gln Arg Gln Ala Leu Glu Gln Ala Arg Lys Ala Glu
 330 335 340

GAG GTG CAG GCC TGG GCG CAG CGC AAG GAG CGG GAA GTG CTG CAG CTG
 480
 Glu Val Gln Ala Trp Ala Gln Arg Lys Glu Arg Glu Val Leu Gln Leu
 345 350 355 360

CAG GAA GAG GTG AAA AAC TTC ATC ACC CGA GAG AAC CTG GAG GCA CGG
 528
 Gln Glu Glu Val Lys Asn Phe Ile Thr Arg Glu Asn Leu Glu Ala Arg
 365 370 375

GTG GAA GCA GCA TTG GAC TCC CGG AAG AAC TAC AAC TGG GCC ATC ACC
 576
 Val Glu Ala Ala Leu Asp Ser Arg Lys Asn Tyr Asn Trp Ala Ile Thr
 380 385 390

AGA GAG GGG CTG GTG GTC AGG CCA CAA CGC AGG GAC TCC TAGGGGCCA
625

Arg Glu Gly Leu Val Val Arg Pro Gln Arg Arg Asp Ser
395 400 405

GTAAGGACAG TGCCCGCCAG GGACCATGTA TGTATCATGG CGGAAGAGTT
GGCCCTGACC 685

TGGAATAAAG CAGTTGGTGT TGCTTATGAG GAAGGTTCA G CTTATCCAG
CACAGCCTTC 745

ACGTTTGCC CTCTGCTGTC ACCACTTGGT CAGAAACTTC CAAACGCAGT
GCCCTGTTCT 805

GCCGGTGTGT AAAGCCTCAG CGCACCCAGGA GACCCTAGAG TGGTTCCAT
CTCACAGAGA 865

ATCAGACAGG CCACAGCCCC CTCAGGCAGC CAGGTCACT GAGTATCATT
AAGAGTAGTG 925

ATGGGAAGAT TACAGTCTGA GGGCCAAACG TGCCTGCTTC CTGTTTTGT
AAATAAAGTT 985

TTGTTGGAAC ACA 998

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 205 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Leu Arg Ala Leu Ser Arg Leu Gly Ala Gly Thr Pro Cys Arg Pro
1 5 10 15

Arg Ala Pro Leu Val Leu Pro Ala Arg Gly Arg Lys Thr Arg His Asp
20 25 30

Pro Leu Ala Lys Ser Lys Ile Glu Arg Val Asn Met Pro Pro Ala Val

35 40 45

Asp Pro Ala Glu Phe Phe Val Leu Met Glu Arg Tyr Gln His Tyr Arg
 50 55 60

Gln Thr Val Arg Ala Leu Arg Met Glu Phe Val Ser Glu Val Gln Arg
 65 70 75 80

Lys Val His Glu Ala Arg Ala Gly Val Leu Ala Glu Arg Lys Ala Leu
 85 90 95

Lys Asp Ala Ala Glu His Arg Glu Leu Met Ala Trp Asn Gln Ala Glu
 100 105 110

Asn Arg Arg Leu His Glu Leu Arg Ile Ala Arg Leu Arg Gln Glu Glu
 115 120 125

Arg Glu Gln Glu Gln Arg Gln Ala Leu Glu Gln Ala Arg Lys Ala Glu
 130 135 140

Glu Val Gln Ala Trp Ala Gln Arg Lys Glu Arg Glu Val Leu Gln Leu
 145 150 155 160

Gln Glu Glu Val Lys Asn Phe Ile Thr Arg Glu Asn Leu Glu Ala Arg
 165 170 175

Val Glu Ala Ala Leu Asp Ser Arg Lys Asn Tyr Asn Trp Ala Ile Thr
 180 185 190

Arg Glu Gly Leu Val Val Arg Pro Gln Arg Arg Asp Ser
 195 200 205

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 943 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 5..604

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TGTC ATG TTG CGC GCT CTG AAC CGC CTG GCG CAG CGG CCG GGA GAC CGG
49

Met Leu Arg Ala Leu Asn Arg Leu Ala Gln Arg Pro Gly Asp Arg
210 215 220

CCC CCG ACC CCG CTG CTC CTG CCC GTG CGC GGC CGC AAG ACC CGC CAT
97

Pro Pro Thr Pro Leu Leu Pro Val Arg Gly Arg Lys Thr Arg His
225 230 235

GAC CCG CCT GCC AAA TCC AAG GTC GGA CGG GTG CAG ACG CCT CCC GCC
145

Asp Pro Pro Ala Lys Ser Val Gly Arg Val Gln Thr Pro Pro Ala
240 245 250

GTG GAC CCT GCG GAA TTC TTC GTG TTG ACC GAG CGC TAC GGA CAG TAC
193

Val Asp Pro Ala Glu Phe Phe Val Leu Thr Glu Arg Tyr Gly Gln Tyr
255 260 265

CGG GAG ACC GTG CGC GCT CTC AGG CTA GAG TTC ACG TTG GAT GTG CGA
241

Arg Glu Thr Val Arg Ala Leu Arg Leu Glu Phe Thr Leu Asp Val Arg
270 275 280

AGG AAA TTG CAC GAG GCC CGA GCC GGG GTT CTG GCC GAG CGC AAG GCG
289

Arg Lys Leu His Glu Ala Arg Ala Gly Val Leu Ala Glu Arg Lys Ala
285 290 295 300

CAG CAG GCC ATC ACG GAG CAC CGG GAG CTG ATG GCC TGG AAC CGG GAC
337

Gln Gln Ala Ile Thr Glu His Arg Glu Leu Met Ala Trp Asn Arg Asp

305	310	315
GAG AAC CGG CGA ATG CAG GAG CTA CGG ATA GCG AGG TTG CAG CTG GAA		
385		
Glu Asn Arg Arg Met Gln Glu Leu Arg Ile Ala Arg Leu Gln Leu Glu		
320	325	330
GCA CAG GCC CAG GAG GTG CAG AAG GCT GAG GCC CAG CGC CAG AGG GCT		
433		
Ala Gln Ala Gln Glu Val Gln Lys Ala Glu Ala Gln Arg Gln Arg Ala		
335	340	345
CAG GAG GAG CAG GCT TGG GTG CAA CTG AAA GAG CAA GAA GTG CTC AAG		
481		
Gln Glu Glu Gln Ala Trp Val Gln Leu Lys Glu Gln Glu Val Leu Lys		
350	355	360
CTG CAG GAG GAG GCA AAA AAC TTC ATC ACT CGG GAG AAC CTG GAG GCA		
529		
Leu Gln Glu Glu Ala Lys Asn Phe Ile Thr Arg Glu Asn Leu Glu Ala		
365	370	375
380		
CGG ATA GAA GAA GCG TTG GAC TCT CCG AAG AGT TAC AAC TGG GCC GTC		
577		
Arg Ile Glu Glu Ala Leu Asp Ser Pro Lys Ser Tyr Asn Trp Ala Val		
385	390	395
ACC AAA GAA GGG CAG GTG GTC AGG AAC TGAGCACAGA GACTTCTGGG		
624		
Thr Lys Glu Gly Gln Val Val Arg Asn		
400	405	
GGCCCAAATA AGCACAGTGC TTGCCTAGGG TCTGTGTACT GGGATAGGAA		
TTGGTACATC 684		
CCAGGAGGAT GGCTCAGCCG TTTCCAGAGC AACCTCAGTC ACTCCAGGCT		
CGGCACTCAC 744		
CACCTGACTG GGAACCTCCA GATGTCCCTG TTCTGGCACC ACAGTCAAAC		
TGAGGGCAGC 804		
ATTAAACTCT GGGAAAGTTCC TATCGCACAG AGGATCGGAC TGGACTGTGT		
CCCTCTAGAA 864		

GCCAAGCTTG TCTTGTAAGT CTCTTGGAGT CCTGTGAGCC AAATGTTCC
TGCTTTATA 924

AATAAAAGTAT TGGAGCCCA 943

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 200 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Leu Arg Ala Leu Asn Arg Leu Ala Gln Arg Pro Gly Asp Arg Pro
1 5 10 15

Pro Thr Pro Leu Leu Pro Val Arg Gly Arg Lys Thr Arg His Asp
20 25 30

Pro Pro Ala Lys Ser Lys Val Gly Arg Val Gln Thr Pro Pro Ala Val
35 40 45

Asp Pro Ala Glu Phe Phe Val Leu Thr Glu Arg Tyr Gly Gln Tyr Arg
50 55 60

Glu Thr Val Arg Ala Leu Arg Leu Glu Phe Thr Leu Asp Val Arg Arg
65 70 75 80

Lys Leu His Glu Ala Arg Ala Gly Val Leu Ala Glu Arg Lys Ala Gln
85 90 95

Gln Ala Ile Thr Glu His Arg Glu Leu Met Ala Trp Asn Arg Asp Glu
100 105 110

Asn Arg Arg Met Gln Glu Leu Arg Ile Ala Arg Leu Gln Leu Glu Ala
115 120 125

Gln Ala Gln Glu Val Gln Lys Ala Glu Ala Gln Arg Gln Arg Ala Gln
130 135 140

Glu Glu Gln Ala Trp Val Gln Leu Lys Glu Gln Glu Val Leu Lys Leu
145 150 155 160

Gln Glu Glu Ala Lys Asn Phe Ile Thr Arg Glu Asn Leu Glu Ala Arg
165 170 175

Ile Glu Glu Ala Leu Asp Ser Pro Lys Ser Tyr Asn Trp Ala Val Thr
180 185 190

Lys Glu Gly Gln Val Val Arg Asn
195 200

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2852 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: Rat 5'OT-EST-xdel

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION:1026..1270

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION:1799..2235

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:1030..1152

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TGACCTCTGT GGATCTGATA TACATGTAAG TGACAGACCA TCCGAGCTAT
ATAGTGAGAC 60

CTGTGCAAGG AAGGATGGAG TGCACGTTCC CTGATGTTCA GAGCAACCC
GTGTCACTCC 120

AGGTAGGTGA GATGAGAGGA AGAGGGTGGC CTTGGCCTGG GCCTCCTACG
GGCCTGGAAG 180

TTGGGAGAAG GATGTAAGCA GACTCTGTC TCTTCTGAGA AATATCAGGT
ATTGCAGTCA 240

GCCCAGGGCTC CTCAGACCCCT CCTAAGTGCA GATTCTCTGC AGAATCTGGT
GTTGACAACA 300

CTAATGAGTA GGATGAGACT TCAGTTCCCT AGCCCTCACC GTCAGCTTCT
GATTACCAAC 360

AACTCTCCA GAGGAGAGCC ATCTACCTTT GGGACAGATG CTCTCTGCC
TGCACTGCCT 420

CCTGTTCTC TTCATTGTAG AGGAAGATAG TACTTTAAAA GCTTCATAAAA
TGGTCTCAAG 480

GTGGGAAGAC CCCGGCTCAG GTGAAAGAGG ACAAGCGTCA CCTCACACAG
GCCACCCAGT 540

AGAAAACAAG TGATCACTGA TACTGAGAAC TCTGGCAATT GCAGAGCTGC
CCAAGACCAC 600

AACAGGGCAG TGCAATGCAA GGAAAAGGTT TGTTGCTCGA TTGCAAACCT
AAAGTTAAA 660

GTGCATCAGG AGAACGCTTA CTCAGAGAGG AAGTGTAAAGC CTAACCTAAG
TAGCTAGAAG 720

CTCAGAATTCTTGATCAG CCCTGGAAAGG GTACACAGGC CACCGGTGGG
CCAGAGAAC 780

ACACGCTTG GGGCGGTGTC CAAGCTTGTG AACAAAGTAGG CAAGAGCGCC
TGGTGTGTA 840

GCTGTCATTG GCGGGCAATA CAGCCCAGCG AACTGTGGTC TCCAAGGTGC
CCCTCGACCC 900

TCCCACCTCA CCCGAGACTC CAGGGACGCG ATGGGCCAGA CAGCAAGAGC
TCCGCCTACG 960

GGGGCGGGGA CAGGAGATT CC GTGATGCT CCTCGACCAC TTCCGGACAG
GGCGCAGGCG 1020

CTAGCTGTC ATG TTG CGG GCT TTG AAC CGC CTG GCC GCG CGG CCC GGG
1068

Met Leu Arg Ala Leu Asn Arg Leu Ala Ala Arg Pro Gly
205 210

GGC CAG CCC CCA ACC CTG CTC CTT CTG CCC GTG CGC GGC CCA CGG CCC
1116

Gly Gln Pro Pro Thr Leu Leu Leu Pro Val Arg Gly Pro Arg Pro
215 220 225

CGC TCA TTC TCG GCT CCT TTT TCC TCG CAG GAT AGC TAGGTTGCAG 1162
Arg Ser Phe Ser Ala Pro Phe Ser Ser Gln Asp Ser
230 235 240

CTCGAACAC AGGCCAGGA GCTCGGCAG GCTGAGGTCC AGGCCAGAG
GGCCAGGAG 1222

GAGCAGGCTT GGGTGCAACT GAAAGAACAA GAAGTTCTCA AACTGCAGGT
GGGCCGAGGT 1282

CGTGAGGAAT GTGGGTATTG GAGATTCCGG TGAGGGAGGC TCTGGGGAGA
GCAGCACAGG 1342

GTGTCAAGTG ACCAGTCTTC AGGAGGCTTC TCTCTCTGCT CTGCACACAC
AGAGTGCCTC 1402

CCAGACAATG GTCAATGAAA GGTTACAGGC TAGTATTGCC GTGTGAAACT
TGAAGGTCAG 1462

GGAAACCATA AATGAGAATG GAGCTGTTT TATTGTGAA GGGAGAGTGA
CAAGGTTGAG 1522

AGAGTCCACC ACCCCGCACC TCCCCCGCC CCCAATCAGG TTGTCACGAT
TCGATTGTT 1582

CTTGGGTTGT GGCTGAGAGA TCTGATGGGT AATTGTCCGA GGAAGAGGGA
TATAATGGTT 1642

GAGGTCACCT AGTACAGTTG TGCTGGCCTA TTGGTGGGAC ACTCAAAGGG
GCCCTGGGCT 1702

CTTTGACAC CCTTCTTAAG GTGGGCTAGA GACAGTAAGT TATGCAGGCA
GCCAGCTCTG 1762

AGAGATCCCA CGTAGCTAAC CTTCTCTTC CCGTAGGAGG AGGCCAAAAA
CTTCATCACT 1822

CGGGAGAACC TGGAGGCACG GATAGAAGAG GCCTTGGACT CTCCGAAGAG
TTATAACTGG 1882

GCGGTACCCA AAGAAGGGCA GGTGGTCAGG AACTGAGAAC AGAGGCCTCT
CAGGCCAAA 1942

TAAGGACAGT GCTTGCCTAG GGACTGGATA TTGGGGTAGA AATTGGTGCA
TCCCAGGAGG 2002

GTGGCACAGC CTTGTCCAGA GCAGCCCCA TTCATTCTAG ATTTGGCACC
AGGTATAGTA 2062

CCTGTTCTGA CACCACATAC AAACCTCCGA CAGCATTAAA CTCTGGGAAG
TTCCTATCAC 2122

ACAGAAGATC AGACTGGACT GTCCCCTCTA GAAGCCAAGA GCTGTCTCCT
GAGTTTCTTG 2182

GAATAGTGTG AGCCAATGT TTCCTGCTTT TATAAATAAA CTATTGGAAA
GCAAAGCCTT 2242

TTGTTATGTG GCTTGCTTT TCTTGTGTA GAATAAGTTT ATTTGTCCCA
GTTATTTGGG 2302

TCTTAAGGTT ATTAGCCAAA AGCCAGTTCA CCTAACTGAG CCAGGAGTTA
GTTATCTGCT 2362

TTGCTCAATC CTGGGCTTG CTGGGTAGGG TCAGGTGTGT CCAAGGTCCA
GAAAGCCAAA 2422

AGGGTGCCCC GTTTCTCCTG GGAAGGCTTC CCCGTCAGTG ATTTCTGTAA

CCGGACCCTG 2482

CCCTGACACA GCGTCATTGG ACTACCCAGC AGACAGTAGA CTCCACTCTA
AACCCGCTTC 2542

TTGCGGTCAAG TTGCTGTCCT TCAGTGTGTG TAAGCAGTGG CCAGACAGCA
CCCTTGGGTG 2602

TCATTTCAAG ACTCTCTCAC CTTGGTCTGC TTTACGTTG GTTGATTTG
GTTTGTCTG 2662

GTTTTGAGA CGAGGCCTT CACTGGAACC TGGCACTCAG TATTAGACT
GCCCAGCCAG 2722

CTAGCCTCAG AGAATGCATC TGCATGCT TGCCTGGCGC TGGAATTCGG
TGCACATGGC 2782

TTTGATGTGT ACCGGGGATC AGACACAGAT GTTTCATGAG TGCAGTGCAT
GCCTGTTAGT 2842

GGTAGAGCTC 2852

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met Leu Arg Ala Leu Asn Arg Leu Ala Ala Arg Pro Gly Gly Gln Pro
1 5 10 15

Pro Thr Leu Leu Leu Leu Pro Val Arg Gly Pro Arg Pro Arg Ser Phe
20 25 30

Ser Ala Pro Phe Ser Ser Gln Asp Ser
35 40

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TTCACACCAAC TCTGTCGAAC

20

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AGGAGGAAGA CAGGTGAAAG

20

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TCATGTTGCG GGCTTGAAAC

20

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TCTTTCAGTT GCACCCAAGC

20

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GTGATAGGAA CTTCCCAGAG

20

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GCCTCGTGCA ATTTCCCTCG CACCTCCAAT GTGAACCTCTC GC

42

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TCCTGCGAGG AAAAAGGAGC CGAGAATGAG CGGGGCCGTG GG

42

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3264 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

- (B) CLONE: Rat 5'OT-EST

(ix) FEATURE:

- (A) NAME/KEY: exon w
- (B) LOCATION:1026..1241

(ix) FEATURE:

(A) NAME/KEY: exon x
(B) LOCATION:1332..1478

(ix) FEATURE:

(A) NAME/KEY: exon y
(B) LOCATION:1559..1682

(ix) FEATURE:

(A) NAME/KEY: exon z
(B) LOCATION:2211..2647

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TGACCTCTGT GGATCTGATA TACATGTAAG TGACAGACCA TCCGAGCTAT
ATAGTGAGAC 60

CTGTGCAAGG AAGGATGGAG TGCACGTTCC CTGATGTTCA GAGCAACCC
GTGTCACTCC 120

AGGTAGGTGA GATGAGAGGA AGAGGGTGGC CTTGGCCTGG GCCTCCTACG
GGCCTGGAAG 180

TTGGGAGAAG GATGTAAGCA GACTCTGTT TCTTCTGAGA AATATCAGGT
ATTGCAGTCA 240

GCCCAGGCTC CTCAGACCCCT CCTAAGTGCA GATTCTCTGC AGAATCTGGT
GTTGACAACA 300

CTAATGAGTA GGATGAGACT TCAGTTCCCT AGCCCTCACC GTCAGCTTCT
GATTACCAAC 360

AACTCTCCCA GAGGAGAGCC ATCTACCTTT GGGACAGATG CTCTCTGCC
TGCAGTGCCT 420

CCTGTTCTC TTCATTGTAG AGGAAGATAG TACTTTAAAA GCTTCATAAAA
TGGTCTCAAG 480

GTGGGAAGAC CCCGGCTCAG GTGAAAGAGG ACAAGCGTCA CCTCACACAG
GCCACCCAGT 540

AGAAAACAAG TGATCACTGA TACTGAGAAC TCTGGCAATT GCAGAGCTGC
CCAAGACCAC 600

AACAGGGCAG TGCAATGCAA GGAAAAGGTT TGTTGCTCGA TTGCAAACCT
AAAGTTAAA 660

GTGCATCAGG AGAACGCTTA CTCAAAGAGG AAGTGTAAAGC CTAACCTAAG
TAGCTAGAAG 720

CTCAGAATT CTTGCATCAG CCCTGGAAGG GTACACAGGC CACCGGTGGG
CCAGAGAACCC 780

ACACGCTTG GGGCGGTGTC CAAGCTTGTG ACAAGTAGG CAAGAGCGCC
TGGTGTGTA 840

GCTGTCATTG GCGGGCAATA CAGCCCAGCG AACTGTGGTC TCCAAGGTGC
CCCTCGACCC 900

TCCCACCTCA CCCGAGACTC CAGGGACGCG ATGGGCCAGA CAGCAAGAGC
TCCGCCTACG 960

GGGGCGGGGA CAGGAGATT CCGTGATGCT CCTCGACCAC TTCCGGACAG
GGCGCAGGCG 1020

CTAGCTGTCA TGTTGCGGGC TTTGAACCGC CTGGCCGCGC GGCCCAGGGGG
CCAGCCCCA 1080

ACCCTGCTCC TTCTGCCCGT GCGCGGCCGC AAGACCCGCC ACGATCCGCC
TGCCAAGTCC 1140

AAGGTGGGC GCGTAAAAT GCCTCCTGCA GTGGACCCTG CGGAATTGTT
CGTGTGACC 1200

GAGCGCTACC GACAGTACCG GGAGACGGTG CGCGCTCTCA GGTGTGTGTA
AAGGGCAGGC 1260

GGCCTCGGC GCCCCCTGGG AAGTGCTGGG GCTGGAGGAT GGGTGCTCAC
TTGAAGCCCG 1320

TCCTCACCCA GGCGAGAGTT CACATTGGAG GTGCGAGGGA AATTGCACGA
GGCCCGAGCC 1380

GGGGTTCTGG CTGAGCGCAA GGCGCAAGAG GCCATCAGAG AGCACCAGGA
GCTGATGGCC 1440

TGGAACCGGG AGGAGAACCG GAGACTGCAG GAACTACGGT GCGAGAGGCG

CGGGGCTGGG 1500

TGGGCTGGGC TAGGCTCACC CACGGCCCCG CTCATTCTCG GCTCCTTTT
CCTCGCAGGA 1560

TAGCTAGGTT GCAGCTCGAA GCACAGGCC AGGAGCTGCG GCAGGCTGAG
GTCCAGGCC 1620

AGAGGGCCA GGAGGAGCAG GCTTGGGTGC AACTGAAAGA ACAAGAAGTT
CTCAAACACTGC 1680

AGGTGGGCCG AGGTCGTGAG GAATGTGGGT ATTGGAGATT CCGGTGAGGG
AGGCTCTGGG 1740

GAGAGCAGCA CAGGGTGTCA AGTGACCAAGT CTTCAGGAGG CTTCTCTCTC
TGCTCTGCAC 1800

ACACAGAGTG CCTCCCAGAC AATGGTCAAT GAAAGGTTAC AGGCTAGTAT
TGCCGTGTGA 1860

AACTTGAAGG TCAGGGAAAC CATAAATGAG AATGGAGCTG TTTTATTGT
GTAAGGGAGA 1920

GTGACAAGGT TGAGAGAGTC CACCACCCCG CACCTCCCC CGCCCCAAT
CAGGTTGTCA 1980

CGATTGATT CGTTCTGGG TTGTGGCTGA GAGATCTGAT GGGTAATTGT
CCGAGGAAGA 2040

GGGATATAAT GGTTGAGGTC ACCTAGTACA GTTGTGCTGG CCTATTGGT
GGACACTCAA 2100

AGGGGCCCTG GGCTCTTTG ACACCCTCT TAAGGTGGC TAGAGACAGT
AAGTTATGCA 2160

GGCAGCCAGC TCTGAGAGAT CCCACGTAGC TAACCTTCT CTTCCCGTAG
GAGGAGGCCA 2220

AAAACCTCAT CACTCGGGAG AACCTGGAGG CACGGATAGA AGAGGCCTTG
GACTCTCCGA 2280

AGAGTTATAA CTGGGCGGTC ACCAAAGAAG GGCAGGTGGT CAGGAACGTGA
GAACAGAGGC 2340

CTCTCAGGCC CAAATAAGGA CAGTGCTTGC CTAGGGACTG GATATTGGGG
TAGAAATTGG 2400

TGCATCCCAG GAGGGTGGCA CAGCCTGTC CAGAGCAGCC CCCATTCA
CTAGATTGG 2460

CACCAGGTAT AGTACCTGTT CTGACACCCAC ATACAAACTC CGGACAGCAT
TAAACTCTGG 2520

GAAGTTCTA TCACACAGAA GATCAGACTG GACTGTCCCC TCTAGAAGCC
AAGAGCTGTC 2580

TCCTGAGTTT CTTGGAATAG TGTGAGCCA ATGTTTCCTG CTTTTATAAA
TAAACTATTG 2640

GAAAGCAAAG CCTTTGTTA TGTGGCTTGC TTTTCTTGT TGTAGAATAA
GTTTATTGT 2700

CCCAGTTATT TGGGTCTTAA GGTTATTAGC CAAAAGCCAG TTCACCTAAC
TGAGCCAGGA 2760

GTTAGTTATC TGCTTGCTC AATCCTGGC TTTGCTGGT AGGGTCAGGT
GTGTCCAAGG 2820

TCCAGAAAGC AAAAAGGGTG CCCCGTTCT CCTGGGAAGG CTTCCCCGTC
AGTGATTCT 2880

GTAACCGGAC CCTGCCCTGA CACAGCGTCA TTGGACTACC CAGCAGACAG
TAGACTCCAC 2940

TCTAAACCCG CTTCTGCGG TCAGTTGCTG TCCTTCAGTG TGTGTAAGCA
GTGGCCAGAC 3000

AGCACCCCTG GGTGTCATT CAAGACTCTC TCACCTGGT CTGCTTACG
TTTGGTTGA 3060

TTTGGTTGT TCTGGTTTT GAGACGAGGC CTTCACTGG AACCTGGCAC
TCAGTATTAA 3120

GACTGCCAG CCAGCTAGCC TCAGAGAATG CATCTCGTA TGCTGCCTG
GCGCTGGAAT 3180

TCGGTGCACA TGGCTTGAT GTGTACCGGG GATCAGACAC AGATGTTCA

TGAGTGCAGT 3240

GCATGCCTGT TAGTGGTAGA GCTC

3264

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44576 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Cosmid DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CGGGCCGCAT AATACGACTC ACTATAGGGA TCTGGTGGAG GACCTATGGC
CCCGAGCTA 60

GAGAAGTGGT TCTAACCTT CCTAGTGCTG AGACCCTTA ACACAGTTCC
TCGTGTTGTG 120

GGGAAACCCC CTCCTGCAAC CATAAAATAA TTTTGTTAC TACTCATAA
CAAGTGTGTC 180

TACTCTATTG CTATGAATTG TAAAATAAT GTGTCTTCCA ATGGTCTTAG
ATGACTCCCG 240

TGAAAGGGTC ATTCTACCCC TAAGAGGTCA TGATCTACAG GTTGAGAAC
ACTGATCTCC 300

AGTAACCTTC ACTTGAGTCC ATATCCTCCA TGAAGGTATG GAAGTCAATA

AAACTGAGCT 360

TCAAGCCTCA TCAAAATGGG TCCATCCCCT GGTACAGTGT GAGTGGAAAGA
ATACCCACCA 420

TACGGTCACT GGAAGGAGGA TGTCTGAAGG GTCTTAGATT GTGTCAAGGG
GTCCTGGGTG 480

TCAGGATCTG ACGAAGCAGG CTCGTCATGT TTCATGAAGA CTACAGGTAT
GTGATAAAAC 540

TGCAAGCTGG AAAAGTACCC ACTGAGCCCG TGTGGCTCTG CTGGGATTTG
GAGGCATGAG 600

GAGCAGAGGG TCTGGAGGAC AGCAGTCCC GAAATAATCT ATGACTAAGA
AGGCTGAAC 660

GGGGTGACTC TCTGGTGGAA AGAGTTGCCT TTTAAGAAGG AAGACATACC
AGGCATAGCA 720

ACAACTGCCT TTAGTACTAG CACTCTGAAG GCAGAGGAAG TCCGATTCT
CTGAGTTCCA 780

AGCCAGCTTG GTTTACACAG CAAGTTCTAG GCCAACTAGG GTTACATAGT
GGACTCTCCT 840

CAAACGGGGT TGAGAAAGGA CTCAGCAGTT AGCTCAGTTA ACTCCAGTTC
TAGGAAATAT 900

GATCCCTTAG TCTGACCTCT TGGCATGTAA GTGGTGCACA TACATATATG
CACACAAAAT 960

ACATCAATCT GCAAAGGGGG AGGGAGGAAG GGCTGGAGTC TGAAGAAATA
GTTCAGTGGT 1020

TAAGAGAATT CACTGCTCTT CCCAATAGCC AAATTCAAGCT CCTAGCATCC
ATGTCAGATG 1080

GCCCACGAAC ACCTGTAATT CTAGCCCCTA AACTCAGTGC CCCTTCACAA
GACGGGGACA 1140

CACGTACACA TATACCTAAA AAATTAGGTG GTTTTTTTT ATTTATAAGG
TCAAATGCAG 1200

AATATCAAAT GGGTAGACA GCAGCTCAA GCTGGCCTCT TCCTCCCAGG
GCTCTTCTT 1260

ACTCTGGCA CCCTCTTGG GTCCAGAAC CAGACATTAG CCATGACTCA
GCTGATAAAA 1320

TGCAACCCAT GGCTCATTAA TTAGGAAGTC TGTAATTAGC CTGTCTGGTA
GCCTCCAGAG 1380

AGAACCCCTT TCACCTGTCT TCCTCCTCTC ACCCAGGGGA AGAGCTCAGT
TTTGGCCCTG 1440

AGACAGAAGA AGGGAACGAG ACCATGAGCA ACGGGAAATG AGATGCTGGC
GCACACACAC 1500

TTTATGTGTG TGAAGTCTCA GAGAGGTCAC CAATAATGAG GCAATGGAAA
TGAGCTGAGC 1560

TGCCTGAACC TCCAAGTTTC CTCCAAGAAA ACCCCACAGG GGAGATGGGG
CATGGCCCAG 1620

GCCAGCTGCC CCAGCCTCTG CTGGCAGAAA GTGAGCCCGC TGCCATTAA
ATTTTGATA 1680

CAGGGTCTCA CTCTACAGCT CTGGGGGCCT AAAACTCACT ATGTAGACTT
CAAACCTAAC 1740

CAAACCAACA ACAAAAACAA ACAAAACCCC TGCAGTGACT GGAGAGATGG
CTCGGTTGAG 1800

AACAATGGCT GCTAGGAGTC AAACCCAGGT CCTGTGGAAG AGCATGCTGG
TAACTGCTGG 1860

GTCATCGCTG GGTCACTCTC TTCACACACA CACACACACA CACACACACA
CACGGCAATG 1920

AACTCTTCAG TGTCTTGATT TACGGTTCT TCCGATAAAAT CCTCAGGAGG
GCAGTCAAGT 1980

GGCTCATTG GCAAATGCTT GCCTGAGACC TGAGTTGGT TCCCAGAAC
CATGGAGGCA 2040

GAAGGAAAGG GCTCCACAAA GCTCTTTCT GAACTCCATA TGTGCACACA

CACCCACTTC 2100

GCACACATTCAATAATAGTGA TGAATGAAAAA TGAAGACAGA TAAAAAAAAC
CAATTCGTG 2160

AAACTGTTAG CACGTTCACT CAATGGCTT GGGGGTAACC TGTTTCAGAG
CCATGGTACT 2220

CAGTCACTAG GCTCATACTG GTCAGACGCT GAGGTCAGCA ATGGAGAGCT
GCTACACCTA 2280

AAGGTAGCAG AGGTCAATTG GCTCTGACTC AGAATATTCC AGCTCTCCAC
ATTCACAGAA 2340

GTTCTACTTG GTCGTAGAAA AAAGCTGAGC CTTTTTTTT TTTTGGAACT
TTATTTTTT 2400

AAAGATATAT TTATTTATG TATATGAGTG CACTGTAGCT GTCTTCAGAC
ACACCAGAAG 2460

GGGGCATCGG ATCCCATTAC AGATGGCTGT GAGCCAACAT GTGGTCGCTG
GGGATTGAAC 2520

TTAGGACCTC TGGAAGAGCA GTCAGTGCTC TTAACCGCTG AGCCATCTCT
CCAGCCCTGG 2580

AACTTTATTT TGAACATGCA ACCCCACCTA CCACTATGGG TTCAGTCACC
AGCGCCTTAG 2640

GAATAAAATT GGAGAAAATA AGCTTATGG TTAGTCAGCT GTCAGCTGTG
GGGTTGGGGA 2700

CAGAAGAATG GTTATGTTT GTTTCCCAT CAAGGCCTCA CTCTGTGACC
TGGTTGGTCT 2760

GCCACTTGCT CTGTAGATCA GGTTCAATT ACAGAGATCC ACCTGCTCCG
TGTCGCTATG 2820

CTGGGATAAG AACTAAGTCA CCCTGCCTAC CTTATTACTT TGTATTCTTG
GGCATGGAAC 2880

TCAATTCCCTT GTCAGCGAGA GAATAACTTC CTCGATCGGA GTGTTTTAT
GTGAATTGGG 2940

CCAAAAAGAC TGCGATGCTC TGAGACCTAT TTGTGAAGCC AAGAGTAGTG
GGTAGCACAA 3000

GTCAGAAATC CAAGGACTTG GTAAGCTGAG ACAGTAGGAT GTGTGCGCTC
ATGCACACAC 3060

ACACACACAG ACACACACAC AGACACATAC ATGCATGGAC GCACAGAGGC
ACCCACGCAC 3120

ATGTGCCTGG ATGAGGCTTC AGTTCTTCAT AAAGCTGCCT TTGAGTTGT
GCCCTCCCAC 3180

TCTTCCTGAG GACTGGAGTC CTCACACCTT GGGCTGATAG TGCAACACTA
CCTTTTTAG 3240

TGACCTCCTC TTTGCAGTCA CAGGCTGAAG GTACAGGGAG GACTCTAGCG
GCCGTCTGCC 3300

TCTGTTAAC ATGAACCTGC AAGGCAGTGG GCAGCCTCAC CCCTAGCGAT
GGCACTGAGT 3360

GATGCCAGGA ACGCTGTCCT CATGTGCCCT TGGCTGTTGG GGCACAGTGT
GCCTCTGCAG 3420

GGCCAGCCTG ACCGTGTGTG CCAGCCAGAA TGCACAATT CTGCCCGACC
TTGGAAGCTT 3480

TTTGTCTTC CTTGTGAGTT TCTTGTCAACC CAGCAGTGT TCTTGCCTCT
TTGCTTGACG 3540

CCTCTATGGG AAGATGGACA AGACTTTTT TTTTCTACAT CCCCTGCAAA
CAGGTTGTC 3600

ATACCTCTCA GGGGCAGGGG TCTTGTCCCT GTCAAGCGCA GCAGGCCACC
AGACCCAGAA 3660

CTATGAAATC TACCCAACCT GTCTCTGTAC AAAGTTAAC AACAAAAAGA
AACTTGGTTT 3720

TGTTTTGTT TTTTTTTTG TTTTGTGTTTG TTTTGTGTTTG TGAGACAGGG
TTTCTCTATG 3780

TAGTCCTGGC TATTCTGGAA CTTGTTCTAT AGACCAGGCT ATCCTGGAAC

TCAAAGAACG 3840

GCCTGACTCT GTCTCCGGG TGCTGGTCAC TCTGAAGATC TGTGCCACCA
TCATCAGGCT 3900

GGGTTTAAA AGATTATGGT TTATATTAA TGTGTATGAG TGTGTTGTT
GCATGTATAT 3960

CTGTACATGA CAGGTGTGCC TGGTGTAC AGAGGCCAGA AGAACATACC
AGATCCCCCT 4020

GGAAGTGAAG TTACAGACAG TCGTGAGCCA TCTCGGGGTT GCTGGGGACA
GAATCCGAGG 4080

GCTCTTCTTG AGTAGCAAGT GCTTCTAACCC GCTTAGGCCT CTCTGCAGCC
CCCACTTACA 4140

GGATTAAAG GTAGAACAAAG GTTTGTCAACC TGTCCTGGAG ACCCTGGCCT
TTAATTCCAG 4200

AACTCTGGAG GTAGAGACAG ATGATTCTCT ATGAAGTTCA GGCGAGCCTG
GTCTACACAG 4260

AGTGCCGCAT GATAGCAAGA AGAAGATCCT GTTTAAAAA GAGACGAGAG
GGGTTGGGGA 4320

TTTAGCTCAG TGGTAGAGCG CTTGCCTAGC AAGCACAAGG CCCTGGGTC
GGTCCCCAGC 4380

TCCGAAGAAA AAAAAGAGAC GAGAGCCAGT GGTTGGTGCA CGTCTTGAT
CCCAGTACTC 4440

TGGAGGCAGA GGTAGTGGAT CTCTCTTGAG TTCAAGGACA GCGTGGTCTA
CAAAGTGAGT 4500

TTTAGGACAT CCAGGATTAC ACGCACAGAA ACCTTGTCTC ATAAAACAAC
AAACAAGACA 4560

AGACAGAAAC TCTCCTAACG TAGACCGCCA CACCTGATTT TTAAAAGCTC
TCAGTGAAAC 4620

TGAGCATGGT AGCACATGTT TGTAATCCCA GCAGACATGT GGGGAGACAA
AGGAATGGAC 4680

TCAGACTCAG CCGGAGAGCA AGTCACGGC TAGACTGGAC CATTCTACA
ATGAGGTAGG 4740

AATTGGGGTT AGCACATCAA GTAAGTAACC CTGGAAACAA GTTTGACTTG
TCCAAGGTCA 4800

CACAGCAATG TCTGGAAAGC TAAGTCTGGT TCCAAGGCC CCCCTCCCTC
CCTCTCTCCC 4860

TCTCTATAAT TGAAAAGTCC ACTGCTTGGC AAAAACTCCC AGGACTATAT
TAAACACAAA 4920

TGCTGGTGTCTCCTCATGTCT TAGGGCTTT ATCCTAGAAG GAATTCAAAC
ACACAACACG 4980

AATACCCCAC AGAAAGGAGG GCAGGGTGGA GGGGTAAGGG AGAGAGGAGG
AACTTCAGGC 5040

TAATGGGGT ATTAACCAGC TCTGTACCCC ATCCACACAG ACCCAAGTTA
GAAAAGAGCA 5100

GGAGAGGGGG TCTGGAGAGG TTGTTAACTG GCCCAGCAGT TTGGCCTGCT
CTTGCAGGGG 5160

CCCAGCTCTG TTCCCAGCAC CCATTCAGT GGCTCACAAC TTTTAACCTCC
AGCCCCAAGG 5220

ACTCTGCTTC CCTCTGAGAG CTCTGTACTT AACAGGGACA CACAGACACA
TACAATTAAA 5280

AAAATTTTAAAGTGAGAG ACGCTCTAGA CAGGCTAGCA AGTATTGAGT
TGTGGCAGGT 5340

ACAGCTATT TAATAGTGAT TTCAGGTTAG AACCTGGGG AGGGGGAAACC
AGGAGTTAAA 5400

CTATGTTAAA TCAGAAAGAC CCAAAGCCAA TCTGGTGGAA GCTGCCATTG
GAGGTTCTAA 5460

CAAGTCTGGC TTGTCAGGGA AAGGCTCAGA ATGAAGGTTT GAGCTGGGGC
ATCATTAGTG 5520

TATAAAAAGT ATGAAAACAC TCTAGGAAGA AGACAAGAGG AGGAACACCA

CGGAGAGCGA 5580

GCCTTACGAT GTTCCAGCAC GTAGACGCCA AAGTGAAGCC AGGAAACCAA
GCACAGGGAC 5640

CAGGAAAGCC CAAAGTCAT TGTGAAAAGG ACAAAAGGCTT CATCCTGGGA
AACTAGGCTG 5700

GGAGAGGCCG TGTTAAATAA AGACAGACAC ACCCATCAAA ATGACCCACA
GAGGGCTTCA 5760

TGATTACAAT AGTATTCAT AGCGGGATT GGGCAGAAAT CTGAATGCAG
GGGATTACAG 5820

AGTAAATGCT GACTTTGGA TAAGAATGGC AGATCACAGG ACAGGTGTGT
GACTCACATC 5880

TTTAAAGCAC ACTCCCAGGG CAGAGGTAGT GAGTTGAGT TTAGGGGCTA
GTCTGGTCTG 5940

GAUTGGAAAT TCTATGAGAC CCTGTTCTCA AAAACTAAAG TATTGGGAA
AAAAGAACTT 6000

CTGAGGGAAA TGGAGGCCGT TAGGTCTCT CTGGGAGCCC GTGCGGCAGG
TGGCGAGGGA 6060

GGATCTGAAA TGGGGAGAGT CAGCAGACTG CTGGACCTTT CCTAGCCAGC
AGAGATGCTA 6120

AGGCAGGTGA AGATTAGGTC TCATGGACCT GACACCCGTG CACACAGGCA
GCATGGCGCC 6180

TTCAAAGCTC TAGTGGATGT GATTGCCCA GACAAGTCTG CCCCCAAAGCT
CATCTTCGTC 6240

CATTAATAGA AAAAAGGTTT CTTCTGACCA AGGAAGCTGT TCTCTCTGGA
AAACAATCAC 6300

TTAACAAAGGA CATTACTAAC ACGAAGCTGC TGTCCGATCA CATCACCATG
ACGCAAGCAC 6360

TTCCCTGGG GTTCATACGC AGTGAACAG TGCTCACGAC CCTGTGCTAG
GCTTGGCCCT 6420

CACTCCTTT CCGCTGGAAT TAAGTGGGGA GTCAGACACC CCAGAGGACC
TGCCCAAGCC 6480

AGAAAGCTTC AAGCCACAGG AGCCAGTGTG TCCTTGGCTT CCCTACACAT
GAGCTGTCTC 6540

TTATCCTCGA TCGAGGGCCT CACAGTCATT CCTGAAAAGA TCTGGCCCCC
AGCCCTGAGT 6600

ATGGAAGGCT AACTTGGCTA CCAGTCCCCA CTGTCCTTAT TAGGAAGAGG
CAAAACCGTC 6660

CTCTGGCACT CTCTGAAGC ATACTGGTAT ATCCGAGAGA GGTAACAGGA
GCCGATGGGA 6720

GCTGGGAGGG TCCTGGCCTA GGCATAGTCT AGAAGACTTGGCTAAGTAG
TCTGGGTCCC 6780

CAAACCATAA CATTTTCTG GTGACTAAAG AAAAGGAGTC TGTAAGCCTA
AAGCAGAATG 6840

TGGTGATACA CGCCTACAGT CCTAGCACTG GAGAGGTGGA GATAGAAAGA
TCAAGAGTTC 6900

AATGCCAGCT TTCTGCTATG TAGTAAGGTC AAGGTAGCC TGGACTAAC
GACTGCCCTA 6960

GAAACAACCA AATGACTTAC CGTCTAAAGT CAGGAACATAC ACTTGCTTTC
TCAGACTGTG 7020

TCTGTCTGTC TGGGGCTCCT CCCATTCCT CTCCTAACAA CATCCACTTC
CACTCCTGCC 7080

TTAGATCTGA GATAGTACCA GCCTCAGGGC ATGGGGTCTC CCCATAGCTT
TTCCTCTGCA 7140

GTACTGTGGG CTCACCTAGG ACTGTTCTG AACTATATCC TACCCCTAGCT
CTCTACCCCTA 7200

GAAGGCCTGA AACTCACAGA AATTCTCCTG CCTCTGCTTT CCAATGGCTG
GGGTTAAAAG 7260

CATGTGTCAC AACTGTCCTT TTTATTCTT TAATATCGAG ACAGGGTCTC

ACCAAGTTGC 7320

CCCAAGACGC CAGCCACACC TGGGACAGGG CAGGCCTTG GCTCTATGTT
CAGTCTTGAC 7380

TCCATGACTG TGGCCGCTAG CCCATGAGGC TGCACGTGGG AATTCCTTC
TGAAAGCTCA 7440

CCTGGTATCG ATGCTTCCTC TTATCCTACA CCACAACTAA CAAACCTGCC
CCACCTCCTG 7500

GTCCTGACCC TGCTGCAGAC CTGCTAGTCC TTGGTGAATG AGACCTGGGG
ACCCCTCTAG 7560

TCTGTTGAGA GCTGCTGAAA TGCTCAACTA TGATTCCAG GTGACCCCTCA
AGTCGGCTCA 7620

CCTCCCTGAT TGCACACGCAC CAATCACTGT GGCGGTGGCT CCCGTACAC
GGTGGCCAGT 7680

GACAGCCTGA TGGCTGGCTC CCCTCCTCCA CCACCCCTTG CATTGACAGG
CCACACGTGTG 7740

TCCCCAGATG CCTGAATCAC TGCTGACAGC TTGGGACCTG TCAGCTGTGG
GCTCCTGGGG 7800

AGCCACTGGG GAGGGGGTTA GCAGCCACGC TGTCGCCTCC TAGCCAACAC
CTGCAGACAT 7860

AAATAGACAG CCCAGCCCGC TCAGGCAGCA GAGCAGAGCT GCACGACGCG
TCGATCCCAA 7920

GGCCCAACTC CCCGAACAC TCAGGGTCCT GTGGACAGCT CACCTAGCTG
CAATGGCTAC 7980

AGGTAAGCGC CCCTAAAATC CCTTTGGCAC AATGTGTCCT GAGGGGAGAG
GCAGCGACCT 8040

GTAGATGGGA CGGGGGCACT AACCCCTCAGG GTTTGGGGTT CTGAATGTGA
GTATCGCCAT 8100

CTAAGCCCAG TATTGGCCA ATCTCAGAAA GCTCCTGGCT CCCTGGAGGA
TGGAGAGAGA 8160

AAAACAAACA GCTCCTGGAG CAGGGAGAGT GTTGGCCTCT TGCTCTCCGG
CTCCCTCTGT 8220

TGCCCTCTGG TTTCTCCCCA GGCTCCCGGA CGTCCCTGCT CCTGGCTTT
GGCCTGCTCT 8280

GCCTGCCCTG GCTTCAAGAG GGCAGTGCCT TCCCAACCAT TCCCTTATCC
AGGCTTTTG 8340

ACAACGCTAT GCTCCGCGCC CATCGTCTGC ACCAGCTGGC CTTGACACC
TACCAGGAGT 8400

TTGTAAGCTC TTGGGGAATG GGTGCGCATC AGGGGTGGCA GGAAGGGGTG
ACTTTCCCCC 8460

GCTGGAAATA AGAGGAGGAG ACTAAGGAGC TCAGGGTTTT TCCCGACCGC
GAAAATGCAG 8520

GCAGATGAGC ACACGCTGAG CTAGGTTCCC AGAAAAGTAA AATGGGAGCA
GGTCTCAGCT 8580

CAGACCTTGG TGGGCGGTCC TTCTCCTAGG AAGAACGCTA TATCCCAAAG
GAACAGAAGT 8640

ATTCATTCT GCAGAACCCC CAGACCTCCC TCTGTTCTC AGAGTCTATT
CCGACACCCCT 8700

CCAACAGGGA GGAAACACAA CAGAAATCCG TGAGTGGATG CCTTCTCCCC
AGGCGGGGAT 8760

GGGGGAGACC TGTAGTCAGA GCCCCCCGGC AGCACAGCCA ATGCCCGTCC
TTGCCCTGC 8820

AGAACCTAGA GCTGCTCCGC ATCTCCCTGC TGCTCATCCA GTCGTGGCTG
GAGCCCGTGC 8880

AGTTCCCTCAG GAGTGTCTTC GCCAACAGCC TGGTGTACGG CGCCTCTGAC
AGCAACGTCT 8940

ATGACCTCCT AAAGGACCTA GAGGAAGGCA TCCAAACAGCT GATGGGGGTG
AGGGTGGCGC 9000

CAGGGGTCCC CAATCCTGGA GCCCCACTGA CTTTGAGAGA CTGTGTTAGA

GAAACACTGG 9060

CTGCCCTCTT TTTAGCAGTC AGGCCCTGAC CCAAGAGAAC TCACCTTATT
CTTCATTCC 9120

CCTCGTGAAT CCTCCAGGCC TTTCTCTACA CTGAAGGGGA GGGAGGAAAA
TGAATGAATG 9180

AGAAAGGGAG GGAACAGTAC CCAAGCGCTT GCCCTTCCT TCTCTCCTT
CACTTGCAG 9240

AGGCTGGAAG ATGGCAGCCC CCGGACTGGG CAGATCTTCA AGCAGACCTA
CAGCAAGTTC 9300

GACACAAACT CACACAACGA TGACGCACTA CTCAAGAACT ACGGGCTGCT
CTACTGCTTC 9360

AGGAAGGACA TGGACAAGGT CGAGACATTCTCGCATCG TGCACTGCCG
CTCTGTGGAG 9420

GGCAGCTGTG GCTTCTAGCT GCCCGGGTGG CATCCCTGTG ACCCCTCCCC
AGTGCCTCTC 9480

CTGGCCCTGG AAGTTGCCAC TCCAGTGCC ACCAGCCTTG TCCTAATAAAA
ATTAAGTTGC 9540

ATCATTGTGCTAGGT GTCCTCTAT AATGACGCGT CGTGCCCACC
TATGCTCGCC 9600

ATGATGCTCA ACAC TACGCT CTCTGCTTGC TTCCAGGAGCC TGCTGGCCCT
CACCTCTGCC 9660

TGCTACTTCC AGAACTGCC AAGAGGAGGC AAGAGGGCCA CATCCGACAT
GGAGCTGAGA 9720

CAGGTACAC TGTGGTCCGT TCAGGGCTGC TGACAGTGCC GTAGGAAGGG
TCATGGGCTA 9780

GGAGAGAGGG AAACCTTGTC TGAGCAGTCA GACTTTAGGG GAGGTTCCCTG
GAAGGAAGCA 9840

GTTATCTTAT ATGGAGTAGA TGGGTTCCC AGAACGGTAA GAGGGGACCA
GGTGCCAGAG 9900

AAGCCACATA AAGGACAGTG TCCCCAGGCA GGGGATATGC CAGAAAATGA
GAGATACTTA 9960

TCACTGGGCT TGGGATGAGA ACGGGTTAAA CTGGGTACCC TGGCCTCCTC
TGCACAGCTG 10020

GAGGTGGCCG GTGGTATGTT GGCTCACCAAG GACTGGGTAG ATGGTACGAA
ACTGTTCTCG 10080

CCTGAGTACA AAGCCTTCC CACCCAGCTC AAACTCTCTT AGCTCCTTT
TTAGCCAGCT 10140

GCACCGGTTT CTTCCTGTCC ACGGAAGACG GCCATTGCC TGTGTCTGAG
CGGAGTATGT 10200

CCCACATCTA GCCTCAGCCT CGTGCCCAGA TCTGCTGTAC TGTATGTTCA
GCTCTGAGTC 10260

TGCCCTTCCG GCAGGGCTGA AGGGAATCCA GTCACTAGGC TCAAATCTGG
TCAGGTCACA 10320

GGTGGCTCAG TTTGAACAA GCTCGATGGG CAGTAGGCAG TTCACCGAGT
CTGCCTTCCG 10380

TTTGCTGAGT TCCTTGAGG ACTTCCGAGG CACTAGGTGT GTCTTGCACC
CATCAGCCTA 10440

ATTCGGTCCT TGCCACCTTC CTACTAGGGC ATAATAGGTT GGCGGGAGGT
AAAAGCCCAC 10500

CAGCGTGGGG CAGGGGTAAG AGTGAGCGAG CCGTAGGTAC AGGAAAGAGG
ATCTTGAAT 10560

GTGTAGGGCC ATCTGAATGT CGGAGAGGTA AGTCTCTGAG AGACTGCTGC
ACACCGGTGA 10620

CACATCAGAG CTGAGGAGGT CCCCCAAGTG TTGTCTCCCC CGCCCCCGC
CCCATACGAC 10680

TCTGTCAAAG CAGGAGAGGG TTTGAGACC TCATGAGAAC TGATCCTCCT
GATAACCTAG 10740

CCGGTTAGAT TTCCACTCTC GCCCTTACG GCTGCTTCGT CCTAGATAGA

GCCAGAGCAT 10800

CTGGCCGGTG AAGCTGGGAT AGCAGCAGGG TGACCTTAGG TTCCCAACGC
CCCTCTTGGC 10860

CTGGCTCCAG CTGACCCGCG TCCTTCCCCG CAGTGTCTCC CCTGCGGCC
TGGCGGCAAA 10920

GGCGCTGCT TCGGGCCGAG CATCTGCTGC GCGGACGAGC TGGGCTGCTT
CCTGGGCACC 10980

GCCGAGGCAGC TGCGCTGCCA GGAGGAGAAC TACCTGCCCT CGCCCTGCCA
GTCTGGCCAG 11040

AAGCCTTGCAGA GAAGCGGAGG CCGCTGCGCT GCCGCGGGCA TCTGCTGCAG
CGATGGTGCG 11100

CACAAAGCCA GGCGGGCTGA GCATGGGAA TGGATGGGGT GGGTGGGAGG
TAAAGGGGGG 11160

CTAAGTGGGG GACTGAGGAA TCAGGACCGG AGATGGAGGG TGAGTAGTAT
GAAGGGGGTC 11220

GAGAGTTGGA ACGTAGCAGG GTAGGATAAA GGGGATTGTG GGGATGGCGC
CCCTATAGGT 11280

GCGCCCACCC CAGGACGCCT GACCTCACAC AGCCCTTCCT TCAGAGAGCT
GCGTGGCCGA 11340

GCCCCAGTGT CGAGAGGGTT TTTCCGCCT CACCCGCGCT CGGGAGCAGA
GCAACGCCAC 11400

GCAGCTGGAC GGGCCAGCCC GGGAGCTGCT GCTTAGGCTG GTACAGCTGG
CTGGGACACA 11460

AGAGTCCGTG GATTCTGCCA AGCCCCGGGT CTACTGAGCC ATCGCCCCCCC
ACGCCTCCCC 11520

CCTACAGCAT GGAAAATAAA CTTTTAAAAA ATGCACCCCTG GTGTCTGTCT
CTCTTCTGG 11580

GGTGGGGAGA AAAGGGGGGA GAGGAATTGG AGTGGGAAC TTCTACTCTG
CTCTGACTGA 11640

TCCCCACATC CAAAGTCGTG CATAAGATAC GCCCCCACCG CCAGAAGGGG
CAGAACCTAT 11700

AAGTCTTAGA GTATAAAGGA AGCTTCTGCT GCTCCTGGAT ACCCACATAA
TACTCAGAAA 11760

AAAAGGCAAG TCAGAAGAAG GGAAAGATCT GAGATCCAGA GGAGCCTGAA
GGGTCAGGGT 11820

GACTTAGCAA GTTTCTATCT GAGACCGAAA TAAAAGGACA TTGTGGACAA
GAGAAACAGA 11880

GCAGGACATG AGGAGAGACA GGATCAGCAA GAGTGACAGA GAAAGAGGGG
ACAGGCCAGG 11940

GGTGGCCATC TCAGCCCTGA TTTCACCCAG ACTAAGGCAA AAACAACGTG
AAGGACTCTT 12000

AACCAAGGCT GTGCTTGGAT GGGAGGAGAA GGTACAGAGA CATTACCCCA
GACCTAAAGA 12060

AGACAATGCC ACCCGCCTTC TCTCCAGGTG CTCCACCATC AAGACCCAGC
CACTGAGAGG 12120

CAGACTCCAG TAAGAGTCCA GCTACAAGTC CTCTACAGGC ACATGTTCAA
ACCGTCACAC 12180

CCACACTCAG GCAGGGAAAT AGACAAGATA GGCTGGAGTT GTGGCTCAGG
AGCAGAAGTC 12240

TTCCCTAGCT ATGGTTCCAG CACCATGGAG ACGGAAGGGG GACTGAGCGG
GGGGGGGGGC 12300

GGGGGGGAGG AAAAAGTAGC AGCTACTAGG GGCATTCTA TGACCCTTGT
CCTCAACCAT 12360

AGCTAGAGAC CCAGAGGAAC ACAGAAGTCC AGCAGCAAGG CGCACATGCT
TGCAATAGCT 12420

CCCAGACGTA AATACTTCAT TCCGTTGGC ACATCCGGGT CATCAGCACT
TGACTCCCC 12480

CCCCACACTT CTTATTACCT CCTCTTTTT CTAAAATTAGATTTATTTC

ACTTATGTAG 12540

ATGGGTGTTT TTGGTTTGATGAATGTCT GAGCACCATG TGTGTGCCTG
GCGCCTCAGA 12600

GGTCAGGAGA GGGCATCGGG TCCCCTGGAA CTAGTTACAG GTGGTTACAG
CCTACCGTGT 12660

GGGCACTAGG AACTGAATCC CAGTCCTCTT AACTGACCCG CACATCCAAA
CCCAGGCTTC 12720

AGCCCCTCAT CAGCCTGTCC CTCCTCCAGG CCCTCAGGTG TCTCCCGTCT
CCGGCTGCTC 12780

TCCCCAGACAT CCTTCCATCC TCTGGTCTCC CTGCTCCTCG CCCTCCTGTT
AAGATCCTTT 12840

CTCTCTGCCCATCTGTCCTGGGCATCCTCTCCTGCGAGCTGCAGCAAGG
TCAGGATGGT 12900

TTACCTCATT TGGGATGGCC TGCAGGTTCT GAGGTCAGGG GCAACTACAG
AGAAGAGAGA 12960

GATTAGTCTG ATTGACTTAA GGTGGTTCAG CAAGGTCAGC TCTGCCAGA
CTCACGGTCT 13020

TTTACCCAGA TGCCAGCTCT CTTCCCATCT CCTCGGTGCC TATACACCTC
TCTGCATGCC 13080

CCGGTTAGA CAGGTAGCAC AGGGGCCAGG CAGACTCCTA TCCCCAGCCT
CTCCCTTCTGT 13140

GGCCCTCTTA GGGTCTGACC TCCAATAGGG CAGGGCCAGG GAAGGGCCAG
^ACCAAAAGG 13200

GACAGAAAGA AGCGTGGCAG GCGGCATGGG CACACTTGAT TCAACCCCTA
CCCGCTGGTGT 13260

ATGGGCAGCT TTAGAATGAA GGTCAGATTCTCACTTCGAG CCTCTGCGCA
CGTGGAGTGT 13320

TGTAAGCGTC TCGCTTCCT CCACCTGTTT CTGGAAGAAT CAGGCTCCTC
TTCCGCTGGAGG 13380

AGAGAATTAT ACCTGCTCAC CCTACTTCTG CCTACTGGAC ATAATATATA
TTTTTTCTT 13440

TTCAGGGAGT CCTTCCTCA GCTACAGAGC CATTAAAGGG CACTCCCAGA
GTTCACAGCA 13500

GATGCTTGCC TCCTCTCTTC AGCCTCCAGA AGCAGAGAGG CTTGTGAGCA
AATGCCAGGA 13560

CCTCTGACCT CCACACAGAC GCTGTGCTGT GTGCACAGCC CTCAAGCACA
CAGCGAAGCA 13620

ATAGTAAAAA GTAACTTAGA CCATTTCAAG GCTGGGGAGA TGGCTCAGGA
GATAAGAGAT 13680

CACTGCTCAA CTTGAGCCTC GGGACCACAG GTAAGACCAA CTTGTCTGCT
GCAAGAGAGC 13740

TGCCTGGTGA GATTGGGACA CACAGAGGCA GAGTCATCT AGGACCGGGC
ACGTCCGTG 13800

TTTCCGAGG TCCCACACCC GCGGATCCCG GCCCGCAGCA GCTCTCTGCT
CCCAGAACCC 13860

GTGAGAAAGA GACCTCACCG CCTGGTCAGG TGGGCACTCC TGAGGCTGCA
GAGCGGAAGA 13920

GACCACCAAC ACTGCCACC CCTGCCACA TCCCTGGCCC AAGAGGAAAC
TGTATAAGGC 13980

CTCTGGGTTTC CGTGGGGAG GGCCCAGGAG CGTCAGGACC CCTGCCTGAG
ACACCGCCGG 14040

AACCTGAGGG AAACAGACCG GATAAACAGT TCTCTGCACC CAAATCCCAT
GGGAGGGAGA 14100

GCTAACCTT CAGAGAGGCA CACAAGCCTT GGAAACCAGA AGAGACTGCT
CTCTGTACAT 14160

ACATCTCGGA CGCCAGAGGA AAACACCAAA GGCCATCTGG AACCTGGTG
CACTGAAGCT 14220

CCTGGAAGGG GCGGCACAGG TCTTCCTGGT TGCTGCCGCC ACAGAGAGCC

CTTGGGCAGC 14280

ACCCCGCCTG GTGAACCTCAA GACACAGGCC CACAGGAACA GCTGAAGACC
TGCAGAGAGG 14340

AAAAACTACA CGCCCGAAAG CAGAACACTC TGTCCCCATA ACGGACTGAA
AGAGAGGAAA 14400

ACAGGTCTAC AGCACTCCTG ACACACAGGC TTATAGGACA GTCTAACCCAC
TGTCAAGAAAT 14460

AGCAGAACAA AGTAACACTA GAGATAATCT GATGGTGAGA GGCAAGCGCA
GGAACCCAAG 14520

CAACAGAAAC CAAGACTACA TGGCATCATC GGAGCCCAAT TCTCCCACCA
AAACAAACAT 14580

GGAATATCCA AACACACCCAG AAAAGCAAGA TCTAGTTCA AAATCATATT
TGATCATGAT 14640

GCTGCAGGAC TTCAAGAAAG ACGTGAAGAA CTCCCTTAGA GAACAAGTAG
AAGCCTACAG 14700

AGAGGAATCG CAAAAATCCC TGAAAGAATT CCAGGAAAAC ACAATCAAAC
AGTTGAAGGA 14760

ATTAAAAATG GAAATAGAAG CAATCAAGAA AGAACACATG GAAACAACCC
TGGATATAGA 14820

AAACCAAAGG AAGAGACAAG GAGCCGTAGA TACAAGCATC ACCAACAGAA
TACAAGAGAT 14880

GGAAGAGAGA ATCTCAAGAG CAGAAGATT CATAGAAATC ATTGACTCAA
CTGTCAAAGA 14940

TAATGTAAAG CGGAAAAAGC TACTGGTCCA AAACATACAG GAAATCCAGG
ACTCAATGAG 15000

AAGATCAAAC CTAAGGATAA TAGGTATAGA AGAGAGTGAA GACTCCCAGC
TCAAAGGACC 15060

AGTAAATGTC TTCAACAAAA TCATAGAAGA AAACCTCCCT AACCTAAAAA
AAGAGATACC 15120

CATAGGCATA CAAGAACGCT ACAGAACTCC AAATAGATTG GACCAGAAAA
GAAACACCTC 15180

CCGTCACATA ATAGTCAAAAA CACCAAACGC ACAAAATAAA GAAAGAATAT
TAAAAGCAGT 15240

AAGGGAAAAA GGTCAAGTAA CATATAAAGG CAGACCTATC AGAATCACAC
CAGACTTCTC 15300

GCCAGAAACT ATGAAGGCCA GAAGATCCTG GACAGATGTC ATACAGACCC
TAAGAGAAC 15360

CAAATGCCAG CCCAGGTTAC TGTATCCTGC AAAACTCTCA ATTAACATAG
ATGGAGAAC 15420

CAAGATATT CATGACAAAAA CCAAATTAC ACAATATCTT TCTACAAATC
CAGCACTACA 15480

AAGGATAATA AATGGTAAAG CCCAACATAA GGAGGCAAGC TATACCCTAG
AAGAAGCAAG 15540

AAACTAATCG TCTTGGCAAC AAAACAAAGC GAATGAAAGC ACACAAACAT
AACCTCACAT 15600

CCAAATATGA ATATAACGGG AAGCAATAAT CACTATTCT TAATATCTCT
CAACATAAAAT 15660

GGCCTTAACT CCCAATAAA AAGACATAGA TTAACAAACT GGATACGCAA
CGAGGACCC 15720

GCATTCTGCT GCCTACAGGA AACACACCTC AGAGACAAAG ACAGACATTA
CCTCAGAGTG 15780

AAAGGCTGGA AAACAATTT CCAAGCAAAT GGTCAGAAGA AGCAAGCTGG
AGTAGCCATT 15840

CTAATATCAA ATAAAATCAA TTTTTAACTA AAAGTCATCA AAAAAGATAA
GGAAGGACAC 15900

TTCATATTCA TCAAAGGAAA AATCCACCAA GATGAACTCT CAATCCTAAA
TATCTATGCC 15960

CCAAATACAA GGGCACCTAC ATATGTAAAA GAAACCTTAC TAAAGCTCAA

AACACACATT 16020

GCACCTCACA CAATAATAGT GGGAGATTC AACACCCAC TCTCATCAAT
GGACAGATCA 16080

TGGAAACAGA AATTAAACAG AGATGTAGAC AGACTAAGAG AAGTCATGAG
CCAAAGGGAC 16140

TTAACGGATA TTTATAGAAC ATTCTATCCT AAAGCAAAAG GATATACCTT
CTTCTCAGCT 16200

CCTCATGGTA CTTTCTCAA AATTGACCAT ATAATTGGTC AAAAAACGGG
CCTCAACAGG 16260

TACAGAAAGA TAGAAATAAT CCCATGCATG CTATCGGACC ACCACGGCCT
AAAACTGGTC 16320

TTCAATAACA ATCAAGGAAG AATGCCATA TATACTTGGAA AACTGAACAA
TGCTCTACTC 16380

AATGATAACC TGGTCAAGGA AGAAATAAAG AAAGAAATTAA AAAACTTTT
AGAATTAAAT 16440

GAAAATGAAG GTACAACATA CCCAACTTA TGGGACACAA TGAAAGCTGT
GCTAAGAGGA 16500

AAAATCATAG CGCTGAGTGC CTGCAGAAAG AAACAGGAAA GAGCATATGT
CAGCAGCTG 16560

ACAGCACACC TAAAAGCTCT AGAACAAAAA GAAGCAAATA CACCCAGGAG
GAGTAGAAGG 16620

CAGGAAATAA TCAAACCTAG AGCTGAAATC AACCAAGTAG AAACAAAAGG
ACCATAGAAA 16680

GAATCAACAG AACCAAAAGT TGGTTCTTG AGAAAATCAA CAAGATAGAT
AAACCCTTAG 16740

CCAGACTAAT GAGAGGACAC AGAGAGTGTG TCCAAATTAA CAAAATCAGA
AATGAAAAGG 16800

GAGACATAAC AACAGATTCA GAGGAAATTCA AAAAAATCAT CAGATCTTAC
TATAAAAACC 16860

TATATTCAAC AAAACTTGAA AATCTTCAGG AAATGGACAA TTTCTAGAC
AGATAACCAGG 16920

TACCGAAGTT AAATCAGGAA CAGATAAACCC AGTTAACCAA CCCCATAACT
CCTAAGGAAA 16980

TAGAAGCAGT CATTAAAGGT CTCCCAACCA AAAAGAGCCC AGGTCCAGAC
GGGTTTAGTG 17040

CAGAATTCTA TCAAACCTTC ATAGAAGACC TCATACCAAT ATTATCCAAA
CTATTCCACA 17100

AAATTGAAAC AGATGGATCA CTACCGAATA CCTTCTACGA AGCCACAATT
ACTCTTATAC 17160

CTAAAAAAACA CAAAGACACA ACAAAAGAAAG AGAACTTCAG ACCAATTCC
CTTATGAATA 17220

TCGACGCAA AATACTCAAC AAAATTCTGG CAAACCGAAT CCAAGAGCAC
ATCAAAACAA 17280

TCATCCACCA TGACCAAGTA GGCTTCATCC CAGGCATGCA GGGATGGTT
AATATACGGA 17340

AAACCATCAA CGTGATCCAT TATATAAACCA AACTGAAAGA ACAAAACCAC
ATGATCATT 17400

CATTAGACGC TGAGAAAGCA TTTGACAAAAA TTCAACACCC CTTCATGATA
AAAGTCCTGG 17460

AAAGAATTGG AATTCAAGGC CCATACCTGA ACATAGTAAA AGCCATATAC
AGCAAACCAG 17520

TTGCTAACAT TAAACTAAAT GGAGAGAAAC TTGAAGCAAT CCCACTAAAA
TCAGGGACTA 17580

GACAAGGCTG CCCACTCTCT CCCTACTTAT TCAATATAGT TCTTGAAGTT
CTGGCCAGAG 17640

CAATCAGACA ACAAAAGGAG GTCAAGGGGA TACAGATCGG AAAAGAAGAA
GTCAAAATAT 17700

CACTATTGC AGATGATATG ATAGTATATT TAAGTGATCC CAAACATTCC

ACCATGGCTT ACGCTCTAAG ATCAAGAAC GACAAATGGG ATCTCATAAA
ACTGCAAAGC 18660

AACTGTAAGG CAAAGGACAC TGTGGTTAGG ACAAAACGGC AACCAACAGA
TTGGGAAAAT 18720

ATCTTACCA ATCCTACAAC AGATAGAGGC CTTATATCCA AAATATACAA
AGAACTCAAG 18780

AAGTTAGACC GCAGGGAAAC AAATAACCCT ATTAAAAAAAT GGGGTTCAGA
GCTAAACAAA 18840

GAATTCACAG CTGAGGAATG CCAAATGGCT GAGAACACCC TAAAGAAATG
TTCAACATCT 18900

TTAGTCATAA GGGAAATGCA AATCAAAACA ACCGTGAGAT TTCACCTCAC
ACCAAGTGAGA 18960

ATGGCTATGA TCAAAAACTC AGGGGACAAC AGATGCTGGC GAGGATGTGG
AGAAAGAGGA 19020

ACACTCCTCC ATTGTTGGTG GGATTGCAAA CTGGTACAAC CATTCTGGAA
ATCAGTCTGG 19080

AGGTTCCTCA GAAAATTGGA CATTGAAGTG CCTGAGGATC CAGCTATACC
TCTCTGGGC 19140

ATATACCCAA AAGATGCCCC AACATATAAA AAAGACACGT GCTCCACTAT
GTTCATTGCA 19200

GCCTTATTAA TAATAGCCAG AAGCTGGAAA GAACCCAGAT GCCCTTCAAC
AGAGGAATGG 19260

ATACAGAAAAA TGTGGTACAT GTACACAATG GAATATTACT CAGCTATCAA
AAACAAACGAG 19320

TTTATGAAAT TCGTAGGCCT ATGGTTGGAA CTGGAAAATA TCATCCTGAG
TAAGCTAAC 19380

CAATCACAGA AAGACATACA TGGTATGCAC TCATTGATAA GTGGCTATTA
GCCCAAATGC 19440

TTGAATTACC CTAGATACCT AGAACAAATG AACTCAAGA CGGATGATCA

ACCAAGAGAAC 17760

TAATCTAAAGCT GATAGACAAC TTCAGCAAAG TGGCTAGGTA TAAAATTAAC
TCAAATAAAT 17820

CAGTTGCCTT CCTCTATACA AAAGAGAAC AAGCCGAGAA AGAAATTAGG
GAAACGACAC 17880

CCTTCATAAT AGACCCAAAT AATATAAAGT ACCTCGGTGT GACTTTAAC
AAGCAAGTAA 17940

AAGATCTGTA CAATAAGAAC TTCAAGACAC TGAAGAAGGA AATTGAAGAA
GACCTCAGAA 18000

GATGGAAAGA TCTCCCGTGC TCATGGATTG GCAGGATTAA TATAGTAAAA
ATGGCCATT 18060

TACCAAAAGC AATCTACAGA TTCAATGCAA TCCCCATCAA AATACCAATC
CAATTCTCA 18120

AAGAGTTAGA CAGAACAAATT TGCAAATTCA TCTGGAATAA CAAAAAACCC
AGGATAGCTA 18180

AAGCTATCCT CAACAATAAA AGGACTTCAG GGGGAATCAC TATCCCTGAA
CTCAAGCATG 18240

ATTACAGAGC AATAGTGATA AAAACTGCAT GGTATTGGTA CAGAGACAGA
CAGATAGACC 18300

AATGGAATAG AATTGAAGAC CCAGAAATGA ACCCACACAC CTATGGTCAC
TTGATTTTG 18360

ACAAAGGAGC CAAAACCATC AAATGGAAAA AAGATAGCAT TTTCAGCAAA
TGGTGCTAGT 18420

TCAACTGGAG GTCAACATGT AGAAGAATGA AGATCGATCC ATGCTTGTCA
CCCTGTACAA 18480

GCTTAAGTCC AAGTGGATCA AGGACCTCCA CATCAAACCA GACACACTCA
AACTAATAGA 18540

AGAAAAAACTA GGGAAAGCATC TGGAACACAT GGGCACTGGA AAAAATTTC
TAAACAAAAC 18600

AAATGTGAAT 19500

GCTTCACTCC TTCTTTAAAA GGGGAACAAG AATACCCTTC GCAGGGAAAGA
GAGAGGCAAA 19560

GATTAAAACA GAGAATGAAG GAACACCCAT TCAGAGCCTG CCCCACATGT
GGCCCATACA 19620

TATACAGCCA CCCAATTAGA CAAGATGGAT GAAGCAAAGA AGTGCAGACC
GACAGGAGCC 19680

GGATGTAGAT CGCTCCTGAG AGACACAGCC AGAATACAGC AAATACAGAG
GCGAATGCCA 19740

GCAGCAAACC ACTGAACTGA GAATAGGACC CCCGTTGAAG GAATCAGAGA
AAGAACTGGA 19800

AGATCTTGAA GGGGCTCGAG ACCCCATATG TACAACAATG CTAAGCAACC
AGAGCTTCCA 19860

GGGACTAAGC CACTACCTAA AGACTATACA TGGACTGACC CTGGACTCTG
ACCTCATAGG 19920

TAGCAATGAA TATCCTAGTA AGAGCACCAAG TGGAAGGAGA AGCCCTGGGT
CCTGCTAAGA 19980

CTGAACCCCC AGTGAACTAG ACTGGTGGGG GGAGGGCGGC AATGGGGGGA
GGGTTGGGAG 20040

GGGAACACCA TAAGGAAGGG GAGGGGGGAG GGGGATGTTT GCCCGGATAC
CGAAAGGGAA 20100

TAACATCGAA ATGTATATAA GAATACTCAA GTTAATAAAA AAAAAAAA
AAAAAGAGAT 20160

CACTGCTCTT GCAGAGGCC CCAGTTCTGT TCCCCACAAC CTCTTAGGAT
GACTCACAAC 20220

CACCTGTAAC CCCATTTCAG GGGATCTGAT GCCCTTTCT GGTCCCCATG
GGCACTGCAC 20280

TCATTTACAA ATACTTCAC ACAGAGACAC ATGCACAGAA ATGGAAATT
AACAGAATAA 20340

ACATCGGAAC ATTAAAAACA AAACAAAACA AAACAAAAAA CAAAAAAAAC
CCCATAGGAC 20400

TGGAGAGATG ACTCAGTGGT TAAGAGCACT GACTGCTCTT CCAGAGGTCC
TGAGTTAAA 20460

TCCCAGCAAC TACATGGTGG CTCACAACCA TCTGTAATGG TCTCTTCTGG
TGTGCCTGAA 20520

GACAGTGACA GTGTACCCAC ATACATGAAA TAAATAAATC TTTAAAAAAA
AAAAGCCCAG 20580

AAAGTGATGA ACTCTATTAC CACCAAAAG AAAAAAAAGA AAAAGAAAAAA
CTCAAATCAA 20640

TCTTGAAGTC TCTTCGCAT ATCTCTTGG CCTCCACCCT GTCTGTGGAT
CCCACATGTG 20700

GGGGCGGTGG GGCATCTGTG TTCATTGCT GAGTGTGAGA GCCACATAAA
GTGCTGGTTT 20760

ACGTGTTAC TTGTTTCTA GATGGGATGG AGCCCAGGAC CTTAACCTTG
TGGGGCAAGA 20820

CTTGCTCTCC TGAGCTCTAC CCAAGCAGTC TGGATTGCGG GTTCCTGTT
TGTCTGTGAG 20880

CTCTCTGCTT TGTGGTCATT TGTGCCACT GGCTCTAGA TCCCATGACT
TCCCAGAGAA 20940

CGCTGTCCTG CAGAGGCAGA CACAGGGCCC CTGAGCTCAG GCCCGGCCCT
GGAGACAGAA 21000

ACGGAGAGGC CAGTTGATT C TGATATTTC CCGTTGTGGC TTCCTTGGGG
CCTGTGTGAG 21060

AATTCAAGCTC TGTAGAAACC TTATGGTTCT GCAACTACCC TCCCCTGGCC
AAGACCCCTT 21120

CTTCTAGCCC GGGATTACCC CCTAACCTC TGAGGTGCGC GCCAAGGTCT
CCTTCCAGAT 21180

ATGGAAGTGA CTGGATGAGT CCCTGCGGC CTCGCCTGCC TTCCCATCAC

CCAGGGCCCT 21240

GTTCGGCTT GCCCTTCCC ACAGAAGTCC ACCATTGCTG TTTGGACTTC
CAAATGGTGC 21300

TCCTAAGTCT GTCTGCCGCA GGCCTTACCC CAGTCGGGAG TGGGAAACGG
GCCTAACTGG 21360

AGATGACAAC CTGTAGAGAC CCCTCGGTCC TCCTAGCAGC CTGCTGGGCT
GTTCTCCCTC 21420

TGAATTGCCA ATGTCCATGG CGTTCCCGGT GCCTTCCTC CCTCCCGTT
CTGACAATTA 21480

GACGCCAGTC AAGTTGAAA AGGAAATCTG CTTTATTAT TTATTTATGT
TTTTTTTTA 21540

ATTTTTTCT GGTAGTGGCC ATGGGAAACG AAGGAAGCGC CCTAAAGGTA
TCATCACAAA 21600

GCAGGGCTCA GCGGCCGGTC TCAGTGCTGG GAGAAGGCGG CCTCAGGGTC
GCAGGCGGGG 21660

TCCTCGTGGC AGCCGTCTGC ACAGAGAGAC GCCGAGTCAG GGCCGCCCTG
GCCCAGCCCG 21720

CCTGGTCCTG GAGCCCCGGT CCCGTCTCGA CCCCTGCCCG ACTCACCCGG
GCTGCAGCAG 21780

ATGCCGGCGG CGGCGCAGCG GCCCCCGCTC CCGCAGGGCT TCTGGCCGGA
CTGGCAGGGC 21840

GACGGCAGGT AGTTCTCCTC TTGGCAGCGC AGCGCCTCGG CCGTGCCCAC
GAAGCAGCCC 21900

AGCTCGTCCC CGCAGCAGAT GCTGGGCCCG AAGCAGCGC CTTGCCCCC
GGGGCCCGCAG 21960

GGGAGACACT GGTGGGAGGG AAGGGATGAG CGGGGGCGG GAGGGGAGCG
GCCGGGGAGG 22020

GAGACCCTGT GGGGCGGGGG GCTGAGCCGG GCAGGGCGAGG GCAGGGCGAG
GAGCGCGGGGA 22080

GGTGGCGGGT CTCCCTGGCT CTCTCTTGG GCTAAAAGC GGTCGAAGGA
GGCAGTCAA 22140

AAGCTCCTCC GCTCCCTCGA TTCCCAGGCT AGGTGGGCC GGTACGCGGT
CAGCGCGGG 22200

AAGGGGGCGG CGGGGGCGAC CCTGTGGCAG CGGGCCGGGC AGCCCAGA
GCCACGGGTC 22260

GAGGGCGGGG CTCTCACCCT GCGCACGTCG AGGTCCAGCA CCGCGCGTT
GCCGCCAGG 22320

GGGCAGTTCT GAATGTAGCA GGCGGAGGTC AACGCCAGGA GGCGAGCAG
GCAGCAGGCG 22380

AGGCTGGAAC CTGCCATGGC GTTGGTGTTC AGTCCGAGAT CGGTCGACCG
ATCCACCGTC 22440

GGTGATGGTT TCTCCAGCCC AGACCGACCT TTTTATGCCT TGTCCACTGC
CATGGTGGGG 22500

CCCAGTCTAA GAGGGTGACT GCATGACTGG TCACAGCCAG GTCTCTGGG
TCAAACGT 22560

CCACACTGTT TAGAAGCAGG CCCTCATT GCAGGGTCTG GGCTGGGTC
AAGGTACCG 22620

CCTCAGCTAA TGACCTGAGC TCAAAAGGGA CACAGCCTAG AAGGGGAGGC
CTAAGCTACA 22680

AGAGGATAAA GAGACTTGGA GGGGGTAGAG GTGCAGCCTA GCCAAGAGCT
GTTTTTCAT 22740

AGAAATCCAA TACCTCAGAA TGAGGTTGGA TAGCGCAAGT GGGTGAGGAA
GCCCTTACGT 22800

GGATCTAAAG CTTAGATGGG GAAAAGGATC TTGTTCAATC TCTGAGTGCA
GCTCAGCCCT 22860

TCTTCTAACT AGCCCGTAAA ACAAAATATC AGTAGAAATC AAACCCAAAA
ACACAAACAAA 22920

CAGACCAAAA TAAAGTAAAA AGAAAGAAAA ATCACAATAA AAGGAAAAAT

CACACTTGCA 22980

CTTACAACTC TGTATTAGGG CTGGAGAGAT GGCTCAGTGG TTAGGAGCAC
TGACTGCTCT 23040

ACCAAAGGCC CTGAGTTCAA ATCCCAGCAA CCATATGGTG GCTCACAAACC
ATCTGTAATG 23100

GGATCTGATG CCCTCTCTG GTGTGTCTGA AGAGAGCTAC AATGTGCTTA
TATATAATAA 23160

ATAACTTTAA AACTCTTAA AACTCTGTAT TAGAACTTGC TATGAGGACC
AGGCTGGCCT 23220

TGAACTCACA GTGATCTATT TGCTTCTGCC TCCCAAATGC TAGGTACCTA
CACTCCCGTT 23280

TGAGAAAACA CAGGCCATCA GCTGCTTGAG CGTGGCCAAC AGGCGGCCTC
AGCTACAGAG 23340

AGCCATTGT CCTAAGGCCA TACCCTTCCT GGTGGCCACA TGTAATGGTG
GCCCATTTA 23400

GTACATACAA CTAGGCATCT CGTGTGCAT TTCAGGGTTG GGCTGCAGGC
CTGCATAGGT 23460

CTGCATGGGA AGAATGCTAC ATGCAGCTCA GTAGCAGACT GCCTGCCTAG
TGTGTGAGAG 23520

ACCTTGGGTC CAGCCCCAG CATGGTGGTA GCAAGACATT TTGGGAACAG
TTTTGCTTT 23580

AAATTTAAC TTTTATTGT GTGTATGTT GTGTACACAG GTTCCTTGG
CAACCAGAAG 23640

AGGATGTTGT ATCCCCAGGA CGTGTCTTA AAGGGGATCA TGAGTAGGCC
TATGTGGGTG 23700

CTGGGAACAA AATTCAGAAT TCTGCAAGAG CAGTGTGCAG ACTTAACCCT
TAAGCCATCT 23760

CCCCAGCTCC TTGATTTCGC ATTTGAATAC AGTTAAAAT GAAATGCACA
TTAAGCCACG 23820

GTGACAGTGA TGCAAACCTT TAATCCCAGA ACTCAGGAGG CAGAGGCAGG
AGAATGTCTG 23880

TGAGTTCCAG GCCAGCCTGG TCTACAGACC TAGTTCCAGG CCAGCCTGGG
CTACACAAAAA 23940

AAACAAAAGC AAAACCAAAA CAAAATAAAG ACACAGACAA ACCATGGCAG
GAAGACATGG 24000

GAGCCTCAAC CTCTTCATT GACGGCTGAG AAATCGAAAA CAGATGACCA
GGAGAGACCA 24060

AGGTCTCACT GCTGCCTTCA AGGCTTGCCC TCAGTGACTG GAAGATGTT
CACTGGGCCG 24120

CATCTTAATA TCTTACCATC TCAGGGCTGG AGAACTGGCT GAGTGGTTGA
GTTGCTCTTG 24180

CAGAAGTCCT AGGTTTGATT CCGAGGACCC ACAGGGTGGC TGAGATCACT
TATCCCAGTT 24240

CCAGTGGAGC CAGTACCCAA ATAGTGCATT ACACACTTGC AGGCAGAACG
TTCAGACACA 24300

TAAAATAAAA TAAATAGACC TAAAAACATT TAAAAGAAAG GAGAACATT
ATCCAGAGTC 24360

GTTTTATTTT GTTTGAGAT AGACTCTAG TTGACCTGGG ACTGTCTGTG
TAGACTAGGC 24420

TGGGCTTGAA CTCACAGCGA TCCCCCTGCC TCCCAAAGTG CTGGGTGTAC
CACCGTGCCA 24480

GGTACCTAGG CCCCTGTTA AGAAGACACT TGCCATCAGT GGCTGGGTGT
GGTCTTAGCT 24540

GCAGAAAGCC ACCTGGCCCT TCCCAGGTGT CCACATATAA TGGTTGGTCC
ACTTTGGTAC 24600

GAATGCTGGG CACCCCAACT GCATGTCAGC TTTGGGCTTT GGGTTAGCTG
AGGTCTGCAT 24660

ACTGGTTCTA GTTGCCACC CCTTCTCTTC CATAGAGGTG GGGCCTAAGC

CCGTGTTCTA 24720

AACTCCATCT CAGGCTCTCT TAAGAAGTGA CCTGCGACAT CCAGGAAGAA
GTAACAGCCA 24780

GTGCCCCCGA GACCCACTCA CTACATGCAG TCTCAGCCCC TAGAGAGGAT
GGAAAAGCCT 24840

CCGGTCTCCT TGTTCTTATG ATCAGCCTTC TCCTCAAGGA GCTGGGGCCA
GTGGGGCAAA 24900

GCACATTCTC TTCTGACCCCT GAATCACAGA TCCTGAGTCA CTGGTGCAAA
CTATCAAGCG 24960

CTAAGTTGGT GGTGAGGTTG ACCTGTACTA CAAATCACTT CATTCTCAC
CCAGACTAGC 25020

TTATTGGCAT TCCAGGCATA GAAAGCCAAG AGCTTGACCC CCACTATAGC
CCCAGAGAGA 25080

CAGCCCACAT AGTCTGTGGG CATA GTGATC TCATCTTAGG TAATCCATGC
ACATAAATTA 25140

GCATGTCTTG ATAATACATA CCTAATGCTC CTGTTAGGCC AGCATGCCTA
ACATGCTCAC 25200

CAACCCAATC TGTGTTGGG AAAGGCCAAT ATTCCGCAAG GCAGAATGCT
AGTCCTTCAG 25260

GAATGGGGCT GCAGCTGGAC TGGGGAGAAC ACACTGAGGT TATAAGAGGA
CCATTGAGGC 25320

CTAATAGCCA AGGTAGAGTA GGCGGAGCCT TGGGTTACAG TGTCAGCAC
CAGGAGGAAA 25380

GAGTCACTAT CACCATGGGG TTCATCTGTC ACTGGAGGAA GCAGAATATG
AACTAAGAGG 25440

CATATTATGT TGGGTTACGA CTTTAGTTAA GATCTGAGTG TATCCCATGT
GATACATTGT 25500

CAGTCCTTAG GAAGATGTCT TGGGAGATGG TGAGATCTT AAGATAGAGA
CCCAGTGTCA 25560

GGTTCTTAT GTCTCTGACA GCATGCCAT GAAGGAAGTG GTCTCTCCTG
GATCTCTTT 25620

TCAGTTTGC AGGCATGGGA TGAAGGGGTG TATCCTTCCG TGTGCTTCTG
CCATGATGTG 25680

TACTTCAACA TAGACCTGTA AGGAACAGTG GCTACAGATT GTGGACTGAA
GTCTCTGAGA 25740

CTGTGAGTCC AAATAACCCCT TTCTTCTAG GCATGGCGGC ACACACCTGT
AATCTCAGCA 25800

TGCTGGAAAT GTGCAGCAGG ATCAGGAGTT AAAGACCAGT CTCAGATAAA
TGACAGTTCA 25860

AAGCCATCAA GGGGATAATG AGATACTTCC TCAAAAACCA TCAAATTAAA
ACTTTTGTTC 25920

TTATACATTA CAACTTGTCA GGGGTTTGC TATAGTAATT AAAAGTCACC
ACAGGAAACA 25980

AAGGCACGTA AACATAGCAA CATGTGCTAT GTTTAAGGCA ACATGTGCTA
GGAAGGTAGA 26040

TATCACCATG CTGGGTGCTT AGACCAGGGC TATGTCGAGG TCCCGGAGGA
GAGCTGAGGA 26100

AGCCCTGGGT GAATGTATAA TGTATCACGG GCCTCAGACC TGTGAGATCT
GGCAAAGCTT 26160

CCCCCTGCAC GCTGTGGGTG AGGTGAATGG GGATTGGCA GAGCCTTGT
CTGGTCTGAG 26220

TGCAAATGCT GACGGTATGT TCTAGTGGAG GTGTTACAA AGGACGGGCC
AGTGTGCGCT 26280

TTAGCCATAG AAGTGGTGGC TCCCTGATGA ATGTCCACAA CCTGGGATTG
CTGCCACAA 26340

GATCAGCCAG GCCCTCTCCT GCGCTGTGCA GAGTGAACAC ACGGAGGGTTC
TGGGCTGCTC 26400

CAGTGGCTGC TACCATTCTG CCAGAGAGTG CACAGGCCAC CTGACCCAG

CCTTCCTGTC 26460

CATGTGTCTG TCCTTCCTTC ACTCTCTCAC CACCCTGTT AGGGTCCCAG
ATCCAAGTTA 26520

TGTAGGGGGT GGATTAGGAA ATGCTATGGG ATGAGAGGCA GTGTTGGTTG
TCATTCTCCT 26580

TAGGGTAACC TGTGAGTATC AAGGAAAGAA AGTGTACACG CAGAAGGCTC
ACCGTGCTGC 26640

TGCTATGTAC AAGTGAGCAC AAATGTAACC TCTGGAAATA CCCATTATC
ATGTCTGTTT 26700

TGGGGCAGA GCCCAGGCAG GCGTTCTAC TCATGGCCT AGGAGCAGCC
TCTCCTCATC 26760

TGGTATGCAG CCCTCCTTA TCCGAGACGG AGCCTGGTGC CGGGACACAG
GTCATTCCC 26820

TGCAGTTGTA TATTATTGG GCAGCTCACT TCTTAAAAT ATTTTGAAA
AAATTATGTG 26880

TATGAGCCTG CATGTATGTC TGTGAAAAT GTCCACAGAG GCCAGAAGAA
GGTGTCAAC 26940

CCCCTGGAAC TGGGAGTTCC GGGTGGTCGT GTTGGCATA TGGGGCCTGG
AAAATGAACC 27000

CTGGTCTCCT AGAAGAGCAA CCAGTGCCT CAGCTGCTGA GCACCTCTCC
AACTCCTGCT 27060

TCTCTGGACT GGGAGACAAA GGAAAAGTGA GAGACTGATT CTGTTCTGTC
AAGTCTCTGA 27120

GCATAGGGAA GACCTAGGTT CATTCTATGT CATCTGTCTG TCTGTCTGTC
TGTCTGTCTA 27180

TCTATCTATC TATCTATCTA TCTATCTATC TATCTATCTA TCTGAGACAG
GATTCACTA 27240

TGTTAGCCTT GGCTGTCTG GAACTCTATG TAGACAAGGC AGGTCTTAAA
CTCGCAGAAG 27300

ATCCTGGTGG TCTCTCCCCA CTTGCCTGA TTAGGCTCAC TTTAAAGGG
AATGAAATGG 27360

GCTGGGTGTG GCAGTACACA CCTTGCTCTC AGCACTCCGA GGCACAGAAA
GGCAGATCTC 27420

TGAGTTGGG GCCAGCCTGG TCTATGCAGT GAGCTATAAGG CAAGCCAGGG
CTACATGGTA 27480

GGACCTTGTG TTAAAAAGAG CCCAAACAA ATAGCTCACT TGCCCAGGTG
AGGTCCACCA 27540

GCATCTCTAC ATTTGACCG GAAGCTAAGA GGAATCTTA TTACATCACG
CCTGCCACAG 27600

TCTCCATCTT TGTTGCAGCT GGAGTGCTCC CACAGGGCTT CCACTGCACG
CACTGCACCC 27660

GAAGGGGCTT CCACTTCACG CACTTCACCC GAAGGGGCTT CCACTTCACG
CACTTCACCC 27720

GAAGGGGCTT ACACTTGATT CACTTGACCC GAAGGGGCTG ACACTGCTTG
CACTGCACCT 27780

TAAGGGGCTG ACACTGACCC AATGGCACCC GAAGGGGCTG ACACTGACCG
CACTGCACCG 27840

AAGGGGCTGA CATTGCACAC GCTGCACCC AAGGGGCTGA CACTTGCTGC
ACTGCACCC 27900

AAGGGGCTGA CACTTGACCG CACTGCACCT ACCAAGGGTG ACACTGCACC
TGCTGCACCC 27960

AAGGGGGCTG ACACTGCATG CACTGCACCT ACCGGGGCTG ATACTGCACC
CACTGCACCC 28020

AGGGGGGCTG ACACTGCACC CACTGCACCC AGGGGGGCTG ACATTGCACA
TGCTGCACCC 28080

AAAGGGGCTG ACACAGCACC CACTGCACCC GAGGGAGCTG ACACTGCACG
CACTGCACCT 28140

ACCGGGGCTG ACACTGCACC GCTTGTAATG TACATTACTG TTTTTTTTTT

TTCTTTCTT 28200

TTTTCAGAG CTGAGGACCG AACCCAGGGC CTTGCCCTG CTAGGCAAGT
GCTCTACCGC 28260

TGAGCTAAAT CCCCTACCCC TACATTACTG TTTAGAAACA AATTATGGT
CCTTCTCACA 28320

TGCTGCAGGA GATTACACAA AGTTGGGGT TATCAAGAAT GTGGATCACG
GTGGATCATT 28380

TTAGCACTGT CCCCCCCCACA GAAAGGGTCA TTTCTAGACA GAAGAAAATA
GTTTATATGG 28440

AACACTCTG GGCTGGCAG TGGTAGCACA TGCCTTAAAT CCCAGCGCTT
GGGAGGCAGA 28500

AGCAGGGCGGA AGCACGCGGA TGCACGCGGA CGCACGCATA TCTGTGAGGT
TTAGGCCAAC 28560

TCGGTCTATG CAGCAGCTTC CAAGACAGCC AAGGCTGTAT GGAGACCCTG
TCTCGGGGTT 28620

GGTGGGGAAT CTCTCACCG TCTTGGTCAC TTCTTATGT GTGAGACACA
TAGACGTTT 28680

TCTTCTGAAT ATTTTATTGC TGCTTGTGGC ATTACAAACT TAGGGAAAAAA
TTGTTAAATG 28740

CTGCATTCCC AGCACTTGAG CCAGTGAAGT TCAGGCCTCC GCTCGTCTT
TAATGGTATT 28800

TGCACAGGGG ATGCCTTGGC TGAGTGAGTT CTTCCAGAAA ACTCCTGGC
CCTTAACACC 28860

TATTCCAGC ATTTGGAAAT CCGAGGCAGG AGGATTGACA TGAGTTGCAG
ACATAGTCAG 28920

CTAGAAGTGC AGCATTAAAT CCTATCTAA AATAATTATT AGAATAATT
AGGGGGAAAA 28980

GCCTCTAATA GAGATGGGAG AGTGTGCGCA TGACTGCCCT ACTGTGTGCT
TCTAGAAATC 29040

AATATGAATG GGCCAGAACT AGAGAAAAGG CTGTGAGAGG CTGTACCTA
CTGTGTGCAA 29100

CCCACTTCCC TCCTACTATG TGGGTGCTGG GCATGACACG AGGTTATCAG
GCTGGGTGAC 29160

AAGCACCCCT ACCTGTGGC AGTCTGCTG GTCCAACCTA TTTGCATTG
AATCCCAGCT 29220

ACTTCAAACC CCATGGGTGC ATATTACCC ACTTTGGTT TTGGAAACAG
GATCTTAAAA 29280

TAAACAGGTC TCACTCTGTA ACCCATGCTG GCCTGAATT AGCATCTTCA
GCCTCAGTCT 29340

CCCAAGCGCT ACGATTCCT ATGTGCCATA TGTCACAATA CATGCACTTC
AGTTTGTCA 29400

AAAGAAGTGA ACCAGGAATA ACTGGTACCT ACCTATAAGA CTGCTGTGAT
GAAGGAGGAC 29460

ATTGTGTAAA ACGAAACTCA GGATATAGTA AGTGCTAAC ACGTGTAGA
CATGTTGGTC 29520

TCCATGAGGG CACAAACCCA GGGCCTCATG CATGCCAAGA ATTGGCCCTA
TCACTGAGCT 29580

ATACAATTAG TCCCTATGAC CTACTGTGAC CTCAGACGCA CACCATGGAT
CTGACATTGC 29640

ATCAAATCAG AAATGAATT CTGAAAGACT TGCTCATAGC ATGCCCTCCC
ACACCCCCGT 29700

CCCAGCCCCC CCTCTCACTG GCAAGGACAT CTCACTGTGG TGGTGGCAGG
GCCTCTAAAA 29760

CATCATAGGA TAGCTGAGCA GCAGTGGCAC ATGGCCTCTC AGTCCCAGCA
CAGGGGAGGC 29820

AGTCAGCCTG GTCTATAGTG TAAGCTCCGG GACAGCCAGG GCTACATAGA
GAAACCTGT 29880

CAAACCTACC CTACTTAAAA ACAGAAGTAG AGCAGTAGTT TGGATTACAC

TGCTTGACA 29940

CTTGGTGGGT AGCATGTGTG CACCTGCCA GGAGCTATCT GGATTCTCAA
ATGGAAGACA 30000

CAGACACAGA CACAAACACA AACACACACA CACACACACA CACACACACA
CACACACACA 30060

CACACACACA CACACCAGTT AACTTTGAC ACGCCATGAC TAGCTCAAAG
GCTAGGGACT 30120

CCCAAACCTT CCCCTGTCAG CAAATGCTCC CCTCTGGTAC TCCTGAGACT
AAGCTAAGCC 30180

TTCCCCTGCT GTCCCAGGCC CAACGGAGGA AGTGAGCATG GTCACTTACC
TGATTCTTT 30240

TTTTCTTTT TTCGGAGCTG GGGACCAAAC CCAGGGCTTG CGCTTGCTAG
GCAAGCGTTC 30300

TACCATGAGC CAAATCCCCA ACCCCACATA TTCTGATTCT TACATGGCTG
ATTGGCTTTC 30360

TGTCCCTGCA GTTCTTACAT CCTGTCCTTC TTCCCTGAAT CATGAGGACC
CTCTCCTCTC 30420

TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCCCTT
CTCTCTGTGT 30480

CTCTGTGTCT CTGTCTGTCT GTCTGTACA CACACACACA CACACACACA
CACACACACA 30540

CACACTAGCC CATGCAAATC TAAGGGCCCC TTCCCGTCTC CCTTGCGCTG
ACCATTGGCT 30600

CCTGGCATCT TTATTGATCA ATCAAAAACC AATTGGGGAT AAGGACCTAC
AGTGTGTTGGA 30660

CATGAAGATT CCTAATTGG GGGCTGCATT AATTCAAAAC ATTGGAACCA
ATTCCCAACA 30720

ACAACAACAA CAACAAATAA AGCAAAGAAA AAAGTTACA ACGCTGCCTT
CATTATTGA 30780

GAAACAAATG TAAGGAAAAC CATCAGGTAT CTGGACTTTT AAACGGCCGG
GATTTAGAGA 30840

CTCTGGGATG TTTTCTGTGT TGGGGATTGA AGCCAGGGCC CTGGGACATG
GTAAGGAAGC 30900

ACTGTACCAT GAAACTACAC CCCAGAGTCT GATAAGGCTA CTGAATGACA
ATTAAAGATT 30960

CATAATTGCT GAAATTCTGG AAAACTCTAA GCTACCAATT TTGTATATGC
TCAACTTGGT 31020

TTCCTGAAAA CATCTGAGGT TCTTGCACGT AACTTTCCT CAGAGCAAGT
ACAACTAAAT 31080

TCTGACTTTG TGACAATAAA GATTGTCAGG AAAGGCTTG TGAAAATGTT
CAGTCCCCAG 31140

GAGACGTGCC CTCCTGCAGC CTGTGAATGG CGGCCAGGTC ACAAGTCAGC
AGATGCAGTG 31200

GAACGGAGTG TGGTACTTCT GTGAGACACT GCAGGACTGG ATGGATGGCT
TAGTAGTTAA 31260

GAACATGGGC TGCTCTCCA GAGGACCTGG TTTCAATTCC AAGCCTTGGG
CCCTGCAAGA 31320

ACCTTATATA GACTGGCTTC AAGTTCTCCA TGTAGCTGAA TATGACATTA
AACTCCAGAT 31380

ACTCCTGAGT CCTAGGTTA CAGCTGTGA CAGCTATGTT TCTTCCCTGA
CGACCCCGCA 31440

GCCCCCATT TGAGATAGGG ATTTAGGTAG CCCAGGCTGG CCTCACACTG
ACTAAGTGAG 31500

ACTGGCTTA AACTCCTCAT CCTTTAAGGT ACCACCATGA ATTTGCTGTA
TAGCTCTGGC 31560

TGGCCTTATA TAATGTAGAC TAGACTGGCC TTTAACTTTA AAATTGTGCA
CTTTATTTT 31620

TTTTAAATTA TGTATTTAT GTATATGAGT ATACTGTAGC TGTCTTCAGA

CACAGGGCAC 31680

CAGACCTCAT TACAGATGGT TGTGAGCCAC CATGTGGTTG CTGGGAGCTC
AACTCAGGAC 31740

GTCTGGAAGA GCAGTCAGTG TTCTTAACCT CTGAGCCATC TCTCCAGCTC
TCTGCTTAAT 31800

AAGTGCTGGG ATGACTAGCA TGTGTCACCC TCCTGGCCAC TTCTGGTGTGTC
TCCTTCCAG 31860

GCTTTAAAAA ATTATCTGTT GGCATGTCCA CACAGGGTTA TATGCATATG
AACGCAGGTG 31920

CCTGTGGGCT GTCCTGTCCCT GGAACCTGGAC TTACAGATGG CTGTGAGCCA
CTTGATGTGG 31980

GTGCCTGGAA ACTAACTGGG GGTCTGAAAA AGCGGGAAGA ACTCACATGA
CTGTGGAGTC 32040

TGCTACCCCT TTTATTATAA AAGAAAAGAA GATATTAA CAGCACGTAT
GAGACACAAG 32100

TGAAAGCTGT GGCCATGGTC TTCAGGGATG GTTAGGTCCCT GCAAAACTGA
AGGAGGTGGG 32160

CTCTGGGTGT TGGTCACATG GTAGATTGAT AGGCCCTGGG TTCAATCCCC
ACCTCTGCAT 32220

AAAGCAGGCA TGGTGGTTCA CTGCCTGCTT TTAGAGGAAG AGGCAAGAGG
ATTGGTAGAA 32280

TCTCAAGGTC ATTTTCAGCT ACATAGCACC AGTCAAATCT TTGAGTCCAA
GACCAGCCTG 32340

ATCTGTGTAC TGAGTTCCAG GGCAGCTACA TAGTTGAGAC CCTACTTAAA
ATTTCAAACA 32400

ACAAAACCCA CAAGGTTAA AAACTCTATC ACTTTAGTT ATGTTGTGT
GTAAGTGTTC 32460

GTGCCACGG AGGTTAGAGC ACTGTATCCC CCGGAGTGGT GAGCGGGCTG
GCATGAGTGC 32520

TGAGGGCTGA ATTCA GCATC TGTGTTAAC ATATTGTTA AAGCACACGA
AGAGGAAATG 32580

GCCAGTGTCA ACAGGAGCCC AGCCAGGCTT GGGGTGGGAA AATGCTTG
CTTCTATCTG 32640

GCAAGAAAAA AACAAATTCCA AGTTGATCC TTGCCAGACT CTTGGCCTT
TACCAGGCTT 32700

CCTCACAGAG TCTGCTGAA CTGTTCTGC AAATTCGCAG AGGAACCTGA
GATCTCAGGG 32760

CACGTTGGAT ACCCACGTGC TGGAGAAACT GAACAATGAC TTTAGGTTTC
ATCGTGCCTG 32820

GATGAAACAT GAAAATACCC CACACCGCTG AGCTGACAAA TGTGCCTCTC
TCTCTGTAGC 32880

CCTTCAGTCA GCTGAGCAGT TTGCCCTCGC TCGGCTGCAG TACCAGCACA
GGCACCCCCAG 32940

TCTCCCCGAT GAGCGGTCTC ATGAGGATTG AGGTTGGTCT CGAACTCCCT
ATGTAGCTGA 33000

AAATGACCTT GAGATTCTAC CAATCCTCTT GCCTCTTCCT CTAAAAGCAG
GCAGTGAACC 33060

ACCATGCCTG CTTTATGCAG AGGTGGGGAT TGAACCCAGG GCCTATCAAT
CTACCAAGTG 33120

AACAACACCC AGAGCCCACC TCCTTCAGTT TTGTAGGACC TAACCTAGCC
CTGAAGGCCA 33180

TGGCCATGGC TTTCCCTCA GCACCCACTT ATCATGAAGG GGCAAGGGTC
CAGTTCTTG 33240

GTAAAGTATC TACGCTTGTG ACTAGGGAGA TACATCCTGG GCAGGAGTGA
AGGGTTACCC 33300

ATTCAGCAGC AGAGTTCTA GGTTTACTGT GACAACAAAG ATCTAGGAAT
GGCCTAGGTT 33360

GTCCTGACAT GATCCCATT A GCCTACCTCA GATATCTGAA TGCAGGGGCT

CACTGTGTGT 33420

CCCAGTCAGG GACAGTATT ACTACCCTAA AGTGGGTTAC AGCTCTCGGG
GGGGGGGGGC 33480

TGCGTGCAGG ACGACACCTG CACCTTCACA CTTGCTTCTT CAATGGAGTA
AGAGGCTGCT 33540

AACATCCCCA AGGTTCCAT TTCAGCTAGG ATGAGAGTCT GGAGTTCATG
TCCCTGGTAT 33600

TCAAGTATAT GACACTGAAG AGCAAAGAGG CAGAGAGCTC ATCCACTAAC
AGGCATGCAC 33660

TGCACTCATG AACATCTGTG TCTGATCCCC GGTACACATC AAAGCCATGT
GCACCGAATT 33720

CCAGCGCCAG GCAAGCATAAC GCAGATGCAT TCTCTGAGGC TAGCTGGCTG
GGCAGTCTAA 33780

ATACTGAGTG CCAGGTTCCA GTGAAAGGCC TCGTCTCAAA AACCAGAACAA
AACCAAATCA 33840

AACCAAACGT AAAGCAGACC AAGGTGAGAG AGTCTTGAAA TGACACCCAA
GGGTGCTGTC 33900

TGGCCACTGC TTACACACAC TGAAGGACAG CAACTGACCG CAAGAACCGG
GTTTAGAGTG 33960

GAGTCTACTG TCTGCTGGGT AGTCCAATGA CGCTGTGTCA GGGCAGGGTC
CGGTTACAGA 34020

AATCACTGAC GGGGAAGCCT TCCCAGGAGA AACGGGGCAC CCTTTTGCT
TTCTGGACCT 34080

TGGACACACC TGACCCTACC CAGCAAAGCC CAGGATTGAG CAAAGCAGAT
AACTAACTCC 34140

TGGCTCAGTT AGGTGAAGTG GCTTTGGCT AATAACCTTA AGACCCAAAT
AACTGGGACA 34200

AATAAACTTA TTCTACAACA AGAAAAAGCA AGCCACATAA CAAAAGGCTT
TGCTTTCCAA 34260

TAGTTTATTG ATAAAAGCAG GAAACATTGG GCTCACACTA TTCCAAGAAA
CTCAGGAGAC 34320

AGCTCTTGGC TTCTAGAGGG GACAGTCCAG TCTGATCTTC TGTGTGATAG
GAACATTCCA 34380

GAGTTAATG CTGTCCGGAG TTTGTATGTG GTGTCAGAAC AGGTACTATA
CCTGGTGCCA 34440

AATCTAGAAT GAATGGGGGC TGCTCTGGAC AAGGCTGTGC CACCCCTCCTG
GGATGCACCA 34500

ATTCTACCC CAATATCCAG TCCCTAGGCA AGCACTGTCC TTATTTGGGC
CTGAGAGGCC 34560

TCTGTTCTCA GTTCCTGACC ACCTGCCCTT CTTGGTGAC CGCCCAGTTA
TAACTCTTCG 34620

GAGAGTCCAA GGCCTCTTCT ATCCGTGCCT CCAGGTTCTC CCGAGTGATG
AAGTTTTGG 34680

CCTCCTCCTA CGGGAAGAGA AAGGTTAGCT ACGTGGGATC TCTCAGAGCT
GGCTGCCTGC 34740

ATAACTTACT GTCTCTAGCC CACCTTAAGA AGGGTGTCAA AAGAGCCCAG
GGCCCCTTG 34800

AGTGTCCCAC CAATAGGCCA GCACAACGT ACTAGGTGAC CTCAACCATT
ATATCCCTCT 34860

TCCTCGGACA ATTACCCATC AGATCTCTCA GCCACAACCC AAGAACGAAT
CGAATCGTGA 34920

CAACCTGATT GGGGGCGGGG GGAGGTGCGG GGTGGTGGAC TCTCTCAACC
TTGTCACTCT 34980

CCCTTACACA ATAAAAACAG CTCCATTCTC ATTATGGTT TCCCTGACCT
TCAAGTTCA 35040

CACGGCAATA CTAGCCTGTA ACCTTCATT GACCATTGTC TGGGAGGCAC
TCTGTGTGTG 35100

CAGAGCAGAG AGAGAAGCCT CCTGAAGACT GGTCACTTGA CACCCCTGTGC

TGCTCTCCCC 35160

AGAGCCTCCC TCACCGGAAT CTCCAATACC CACATTCTC ACGACCTCGG
CCCACCTGCA 35220

TTTGAGAAC TTCTTGTCTT TTCAGTTGCA CCCAAGCCTG CTCCTCCTGG
GCCCTCTGGG 35280

CCTGGACCTC AGCCTGCCGC AGCTCCTGGG CCTGTGCTTC GAGCTGCAAC
CTAGCTATCC 35340

TGCGAGGAAA AAGGAGCCGA GAATGAGCGG GGCGTGGGC CGCGCACGGG
CAGAAGGAGC 35400

AGGGTTGGGG GCTGGCCCCC GGGCCGCGCG GCCAGGCGGT TCAAAGCCCG
CAACATGACA 35460

GCTAGCGCCT GCGCCCTGTC CGGAAGTGGT CGAGGAGCAT CACGGGAATC
TCCTGTCCCC 35520

GCCCCCGTAG GCGGAGCTCT TGCTGTCTGG CCCATCGCGT CCCTGGAGTC
TCGGGTAGAG 35580

TGGGAGGGTC GAGGGGCACC TTGGAGACCA CAGTCGCTG GGCTGTATTG
CCCGCCAATG 35640

ACAGCTACAA CACCAGGCGC TCTGCCTAC TTGTTCACAA GCTTGGACAC
CGCCCCAAAG 35700

CGTGTGGTTC TCTGGCCCAC CGGTGGCCTG TGTACCCTTC CAGGGCTGAT
GCAAGAAATT 35760

CTGAGCTTCT AGCTACTTAA GTTAGGCTTA CACTCCTCT TTGAGTAAGC
GTTCTCCTGA 35820

TGCACTTTAA ACTTTAGGTT TGCAATCGAG CAACAAACCT TTTCCTTGCA
TTGCACTGCC 35880

CTGTTGTGGT CTTGGGCAGC TCTGCAATTAA ATTGCCAGAG TTCTCAGTAT
CAGTGATCAC 35940

TTGTTTCTA CTGGGTGGCC TGTGTGAGGT GACGCTTGTC CTCTTCACC
TGAGCCGGGG 36000

TCTTCCCACC TTGAGACCAT TTATGAAGCT TTTAAAGTAC TATCTTCCTC
TACAATGAAG 36060

AGAAACAGGA GGCAGTGCAG GGCAGAGAGC ATCTGTCCCA AAGGTAGATG
GCTCTCCTCT 36120

GGGAGAGTTG TTGGTAATCA GAAGCTGACG GTGAGGGCTA GGGAACTGAA
GTCTCATCCT 36180

ACTCATTAGT GTTGTCAACA CCAGATTCTG CAGAGAATCT GCACCTAGGA
GGGTCTGAGG 36240

AGCCTGGGCT GACTGCAATA CCTGATATT CTCAGAAGAG AACAGAGTCT
GCTTACATCC 36300

TTCTCCCAAC TTCCAGGCC GTAGGAGGCC CAGCCAGCAC CCTCTTCCTC
TCATCTCACC 36360

CTACCCTGGA GTGACACAGG GTTGCTCTGA ACATCAGGGA ACGTGGCACT
CCCATCCTT 36420

CTTGCAACAG GTCTCACTAT ATAGCTCCGG ATGGTCTGTC ACTTACAATG
TATATCAGAC 36480

TCACAAGAGG TCCATCTGCC ATTGCCTCCT AAATGCTGGG GTTAAAGGCA
CATACCACCA 36540

CACCTGTCCT AAACCTTCT TCTTCGGGGT CATCCTAGAT AACCAAGTATC
TCATTTCAGA 36600

TAACCTCAGT GTCTGGCAA AGAGAATATT TCTATGGTGT GGGTCATTCC
TAGAGGCTTC 36660

CTAACCTTGC TGGCTCTGAC GTTCTCTCGG CTGGTCAGGT CTACTCATCC
TTCTTCAGA 36720

GGGTTTCATA AGTTGTAAGA GATTTAGGCC TACGGTGGAT GAAAGATGTG
GAGTCATTTC 36780

GAGTAGCTAA TGCTACAGAA CTAGAAGGCA GGTTCTCTGC CCCCTCTCT
GACCTGTTGG 36840

GGAAGTGGAA GTAACTTCC ATTTGTGACC TTCCCCACTA GGTGGCGAGA

TAGAATTGTC 36900

AAAGCTGGGA AAGGAGGCTT TTCTGGCAG TTCATGGGTA GAAGGACAGA
CAGACAGAGT 36960

TGGAGGAATG GAAGCCTCCT CATTACCAA GGGTAAACT ATGGATGAGG
TGACTTCAGG 37020

TGCCTGCAGG ACCCTATGCA GACGGTCCA GGATTAAATG ATCAGGCCAT
TCTATTCCT 37080

CTGGTGTCAA ATCCAGTGAT ATCATTAAAA CAAAAACAAA AAAGCCCCAA
TCAGGGTCTT 37140

ACTTGATGGC CTTATATTC CAACAAAGCC CAGGCTGGCC TTGAACTTGA
AGCAATATCC 37200

CTGCATCTGT CTCCAGAGTG CTAAGATTGT GTGTGTCAAC ATACCAAGGT
ACAGTGATCT 37260

CTTGAAACAG GGAGGTGCAA GTCATTACTC AAACCCCTCC TCACAATGTT
CTATGAGCAA 37320

ATCCGAAGTT GATGTTGGCT TTTAAAGTCA CCAGACAAGT GTCCTCTGC
TTAGATCTTC 37380

CTAGGAACTG AGGTTGAAA CAAAAAGCAT AACATGGTTG GAGAGATGGC
TCAGTAGTGA 37440

AATTCTGAAT GTGGTCCCA GCATCCACAT TGGCACCTC AGAATGGCCT
ATAACTTCAA 37500

TTCTAGGGAC CAAGTAATCT CTTCTGGCTT ATGGGTGTCA CTCACATGCA
TGTGTGCATA 37560

TGGTGCCTAA GTAAAAAAAT AATCTTTAA AAGCAGATT TAAAAAAAAT
TTCAACGATT 37620

TTTTTTAAT GTTCATTGGT GTTTGCCTG CATGTATATC TGTGTGAGGG
TGTCAAGGTTT 37680

CCAGGAACTG GAGTTACAGA CAGATGTGAG CTACCTGTTG GTGCTAGGAA
TTGAACCCAG 37740

GTCCTCTGGA AGAATAATCA GTGCTCTTAC CCACTGAGCC ATCTCTCCAA
CCCAAATACA 37800

TCTTAAAAAA AATTAAAACA GTGGACCTGC CTTCTAGTTC TGTAGCATT
GCTACTCAAA 37860

ATGACCCACA TCTTCATCC ACCGTAGGCC TAAATCTCTT ACACTTATGA
AACCCCTCTGA 37920

AAGAGGATGA GTAGACCTGA CCAGCCGAGA GACATCAGAG CCAGCAAGGT
TAGGAAGCCT 37980

CTAGGAATGA CCCACACCAT AGAAATATTCTCTTGCCTT GACACTGAAG
TTATCTGAAA 38040

TGAGATACTG GTTATCTAGG ATGAACCCCG AGAGAAGAAA GGTTAGGAC
AGGTGTGGTG 38100

GTATGTGCCT TTAACCCAG CATTAGGAG GCAATGGCAG ATGGACCTCT
TGTGAGTCTG 38160

ATATACATTG TAAAGGGGAG AACTCCCGA ATTTGTTCTC TGACCTACAC
ATGTGACATG 38220

CATGTGTCG TGCACACACA CATAACACACA CACACACTGT AAAAATGCAA
AATGGCTACC 38280

AAGTGGTCAT TGAGCTTCTC AACCTCACTG ACAGCTACAT TATTATATAG
ACTTACTGGG 38340

AACAGATCCG CAGGAAATTATTTGGAATCT TTTTCTTTCTC TCTAACGGGG
GCTGATCTGG 38400

AACTTCTGAG CCTTTTGTT CCCTATCATG AATGCTGGGA TGGCAGGGGT
TTCCACATGA 38460

CTCGTTCGAT GTAGTATTGC AGACTGAACG CAGGACTTTC CACACACTAA
GCAGGGCATTC 38520

TGTGAACGT TACGTCTCCA GCCCCATTTC TAAATTCTAA CACCAAAGTG
CTAGTTTGT 38580

CCCTTGACCT GGACACTGCA GTGAGTCAC AGAACTTATA ATCACCCTGT

TTAGTGTAGA 38640

AGCTACCTCA ATCACCATGA CATTTCCTAA AAATGTGTTCACTTCCCTCT
TTAGAGTCCA 38700

AGCACACCAA GCTTGGCGGA ACAATGATAC AGTCTAACTG GATCTGTTTC
AAAATTGCAA 38760

CTTGACTCTA CATCTAAATA GGTATGTGTT GTGACAAGTT TATTATGTTG
TGTGTGTGTG 38820

TACACATGTG CCACAGGAAG CCAAAGGACA ACTTGCTAGA GTGCATTTC
TTCCCTGGGA 38880

ATTGGCCTCT GGTTGTCAGG CTTGGTAGCA CGCACTTGA CCCTCTAAGC
CATCTTGATG 38940

GCCCAGAGAG TGAACCACGC TGTTTCACT TTCCTACTTC TTGGGCTGAA
TTCTCAAGTA 39000

CCTGCCCTTG CAGCTTGCA CCCTCCTAA CTTCAAAAGG AAACTGACAT
GGAGAAGGGT 39060

GATACTTGAG GATTCCTGG CTCACTTAGC TCAGGACTCT GGCTAAGAA
CAGGGAACCC 39120

AGCAGTGTGA ACAGGGTCC AAGAGAGTTC ATTTGTACTT ACCGGCAAAA
CAGTGTGGCA 39180

GGCTTCACAC AAATACATAC TCGGCACCAAG GACAGGGCCA CTCTGGATGG
AGGTGGGCTT 39240

AGGTGGGGTA CTGCCACCC AGGGTTGTCC TCTCTGTAA GCAGACTCAT
GGGGACAGCC 39300

CAGAAGTGAT CCCACAGTCT CTCTGAAGCT GACAATAGGG GATAATTCTA
AGTCCTCATC 39360

CTGTGCTCAT CCACAGTCCT TTGTCGATCT GGACACTACT ATCATGGGCT
GCTGGAAACA 39420

GGTCTTGCA GCCCAAGTCT GAGCCACTAG CTCTGCTTTC ACTGCCAGCC
ATTAACCTCC 39480

GGGAGTGGGC GTGGGATAAG AAGAACATT TATAGAGTCA ACGGCCAATC
TGTATTTGGG 39540

CTGAAAACCA TATTAAGGAA GGGCCAAGCC TGGCATAATG GTGACCAGAG
CCACTAGGGG 39600

ACCAACTGCA CCCAGCTTA GCAAAGTGAC AGGCAGCATG AGGTACCATT
ATGTGTGCTG 39660

GGCATGCGGC TTCAGGATGG CTCTGTGACC TCCTAGAGGT TGTCTTATTG
GCAGGCATAG 39720

GAAACAAAGG CAGAGAATGA ATGCTACAGC CAGAGAGACC CAGATCTGCT
AAGTGGATGA 39780

CTCTTGACA TATGTGTGTA TGTGTTTT GAGGCAGGGT CTCACTGTGT
AGCTCTGACT 39840

GTCCTGGAAT TGGATCTGTT GGTCTCAAGT TCAGATCCTA GTGGTTATT
TTTCCTGTGT 39900

ATGTGTGCTT GTCATGCACA AGCATGTGTT AAGGTGAGTA GATATGTAGG
CACATGGAGA 39960

TCAGAACAAAT GGTGTCACTC CAAACCTTA TAGACCTATA TCCATCTTGA
CATTAGGGTT 40020

ACAGGTGTGT TCAACATAGA TATGGCCAAA ATTTAATGTG GGTTCTGAAG
ATCTAAATAT 40080

GTCTTGCT GGCTAGTTCT ACGTCAACCT GACACAAGCT AGAGTTATCT
GAAGGAAGGG 40140

AACCTTAGTA GAGAAACTGT CTCCATGAGA TCCAGCTGTA TAGCATTTC
TTAATTCTTA 40200

GTTAGAGACT AATGGGGGAG GGCCCAGTCC ATTGTGGGTG ATGCAACCTT
AGACAGGTGA 40260

ACCTGGGTTT TGTAAGAAAG CAGGCTGAGC AAGCCATGAG GAAGCAAGCC
AGTAAGCAGC 40320

ACTGACCATG GCCTCTGCAT CAGCTCCTGC CTCCAGGTT CTGCCCTGTT

TGAGTTCTTG 40380

TCCTGACCTC CTTCCGTGAT GAACAGTGAT ATGGAAGTAT AACCAAATAA
ACCCCTCTCCT 40440

CCCAAGTTGC TTTGGTCATG GTGTTTCATC ACAGCAACAG AAAGCCCAAC
TGAGGCAGGT 40500

TTTCATGCTG TATAACAAGC TGAACCATCT TACCAGCTCC ATAGTGTAA
TTTTAAAAGA 40560

TGAGTGTGTA ACTTTCCCTT TTTTCCTTT AAAAATCCAA AGAACCCACGT
TCCTCAGGAA 40620

AAGCTCTGGG CCAGTTCTCC TGGTAACCTT GAAAGTCTTT TAAAGGCAGA
GTCTATGTTA 40680

GACAAGCTGG CCTCAACCTC ACAGAGATCA CCTCCCTCTG CCCGTTAACT
GCCTGGTGAG 40740

CTACAATGTG TTTTAAAGA TGTCCCTGTT CCCTCTTAAA CAACTCCAAT
TTCACCCATG 40800

TGTTCCCATT TGGTAGGACA GGAAGCCATT TGTTCATCAT GAAGCTTCTG
CTGATGTCAG 40860

GACAGGCGCG CGCGCGCACA CACACACACA CACACACAGC AGCTTAGTC
ATTGTGGTC 40920

AGCTGGGAAA ATGGGAAAAC ACGGTTGGAG CTGAGTTGAA CTGAAGAGTT
GGTGGAGACA 40980

CATGGTGCAA ATCCTGAGCA GTAGCTGAAG GAAAGGTACA AGTTGGCAG
TAGATTGGCC 41040

AATGAGGTGC AGAGATAAAG CAGAAGGGCT GCCCCGAGAG CTGCAGCATG
GTGCGTGGAA 41100

CCCTTCAGGA GGTAGAAAGG TAGAAAGGGC TGCTTGGACT ACTAGTGTGT
AGATTACTGT 41160

CTTTCAGCAG GTGAAAGACA AGGCTAGAGC CTGTGATTGG ACAGTAGAAA
AGGAGGGCGG 41220

GCTGAGAGTT TGAGAGTCTG GAGGGATAGG AGGAAAGAAG GAAGATGGAG
GAAGAGAAGG 41280

ATGACCCAGA GCTGTGTGGC TTTAAATAGC CACAGGTAGC TATGAATATC
ATATAAGGGG 41340

TGGATTATGA CAGGACAATT TGTCCACTCA AGGTGGGCAG CTTATATCAT
ATTAATTGGC 41400

TCTGAGTTCT TTGTCTTGGG CATTGTGA GCTGAGAATT TACTGATATA
AATCTGACTG 41460

ATAAATTACA AGCCTCTAGA GTTTGATT TACTGGGTTA CAGGGATTG
TGACAGTTAA 41520

CTGCGAGATG CTACAGCCAG AGAGACACGG ATTCTGCTAA GTAGATGACT
CTTGTACATA 41580

TGTGTGTATG GGGGCTAGCT GTGAAGGCAG TGAAACTGCT GCCAGGGCCA
GAGAGTAGTT 41640

GGCACTACTG TGGGATGGTG CATCCATTAA TTTAAAAATT ATTAATGCA
ACACTAGTGA 41700

GTCATCCAGT AGGAAATGCT GGGGTCTGGG GAGCTGGGGG TGGAGGAAAG
CCACAAGCCC 41760

ACGGAGCCCC AGATCCCCCA CCTCTTGGA GAATAACACT GATATCAGTG
ACTCAGACAC 41820

AATAGATCTT GGGGTTCAAG ACCCAAGCTC CTCTAGTAAG CATGGGTGCA
AAAGGTGTGG 41880

AATGGAGAGT GAAGGAAGAC TTTTCATAA GCCTGTCACA AATGAGGAGG
AAGCTAAGCT 41940

TGGGAAATGC AGGCCTTCAG TGGCAGACCA AGTGGAGTCA ATGAAGTAAG
GTCTGAGTAG 42000

AAGGGCTCTG GGTGTGCGCT TCAGGCTGGG TGCACACTTC TTTCTGAGGA
AATGCTCACT 42060

TCCACTTGA CCATTCCCTG ACCCAGGTCA TAGCTGATGT GCCAGAGTGT

CATGGGTGAA 42120

GTGGTCACTT TCGCTCTTC CACACAACTG TGCTGTGAA GACACCCTT
CTCTGGTCAT 42180

ACAGGAGTCC CCTGTGGGT TTGAGCCCTG ACTTAAAAAG AAAGGATGAG
GGCTACTTCT 42240

GTGGAAGGGA GCAAAGAGCA GAGGTCAATT CTGCTGGAGG AGATCTGCTA
ACAAGCATGT 42300

GATGTTAAC ATTAAGGGCT GCTCATCAAG TCAGCACTGA CTCCAGCAGA
GTCCTGTCGA 42360

GGCTACTCCA GTATGCCCTG GTCAAGACTA GCCTTGGCAA GGGAGCAGCC
TGGGCTGTTG 42420

CTAGGTGGAT AGAAGCACAC ACAGAGGATT TTCTTAGTGG TGTATGTAAT
CACTGAGGTC 42480

TTGCTGACCC AGTAGGCATA CTCCTCCATT GCTAGACTCA GTCACACAAA
GTGTACAAGA 42540

ACAGGGCATT CTTCATGGAA AATTCTGAC TGGGTCTTT AGAGCTCCAG
TTCCTAGAGG 42600

GGCAGATGAT CCCAGTGAAT TATGCTCAGT GTAAAGCTGG TCTGCTGTCA
CATCTTGCT 42660

CCCAAAGGTT TCTGGGATTCTCCTGTACT TTCTTCTATT TTTATTTCA
AGACAGGGCT 42720

TCTCTGTGTA GTCCTGGCTG CTCTGGAACCTGCTCTGCAG ACCAGGTTGG
CCTCAAATT 42780

ATAGAGATCC ACTTGCCCTCTGCCGTCCAAG TACAGGGATT AAAGTTGCAT
GCCACCACTG 42840

CCCAGCCTCT CTAAAATTTT CTTAATTAAT TTATTTTCA AGACAGAGTC
TCACTATGTA 42900

GTCCTGGATA TGCTGGAACCTAGTAATGTA GAACAGGCTG TCCTTGAACA
TACAGAGTTC 42960

CACCAACCTC TGCTTCCAAG TGCTGGGATT GAAGTGTGTG CCACTATGCC
CAGCTAAAAC 43020

CTGTTTATT TTCTGTGCAT GGGTGTTGC CTGCATGTAT GTCTGTGCAT
CATTGCCTG 43080

ACTGGTGCC CACGGAAAGTCA GAGGAGGACA CTGGATCCCC TGAGGTGCC
ACGGAAAGTCA 43140

GAGGAGGACA CCGGATCCCC TGCAGTGCC ATGGAAGTCA GAGGAGGACA
CCGGATCTCC 43200

TGGATCTGGA TGACTGAGCC ATCACATGGG TCTTGGGAAA AGATCCCGGG
TTTGCTCTAA 43260

GAACAAAGTGC TCTTAATGAT TGAGATGTCT CTCTATCCA TGTTCTTG
TACACAAACA 43320

CCATGGACAC GTGGCATACA CTGGGCTTCC TTTTCACACC ACTCTGTCGA
ACTTAAATTG 43380

TGCTGGCGGC TCCAAGTAC CTTCCCTTC TATTCTAAA TTCTCGGCAT
GGCTTGGGTC 43440

TGGTTAAGTC CCCCCTTTC CAAGCAGCCG GAAGCACTTA TCTCTGAATG
TGCCTCTGTG 43500

GGACACACCG GGGGACCTGC TGAAGCCTCT GAAGAGCAGA GGTGATGTCT
GCCTCCCCAT 43560

CTTGCCTCT TTGTGCTAAG AAGCTACTG TGATGCTGGA GGTGGTGGGG
AAAACCCACC 43620

AGCCTTGCCA CCTGAAGTGA AGGGCAGCCA CGGCCTGTGT CCTAGCCAGT
GGGGATTAGT 43680

GAAAATGGTA AAGTGGCAA CGAGGCTGCT TGCTTCTGA GCTTCCTCCT
ATTTTGGGTT 43740

GGTAGCAGCA GCGGCCAGT TCCTTCCCAC TGTGGGGATG AGGAGTACGC
CCTCAGGATG 43800

CCGGCATCAG AGAAGGCAAG AACAGACGCA GTGTCGCACG TCTTCAATT

CAGCACTTGG 43860

GAGGCAGAGA CAGGCAGATC TCTGCGAGTT CAAGGCCAGT CTGGTCTACA
CAGTGAGTTC 43920

TAGGTTCGTC TGTGTTACAC AGGGAGAACT GTCTGAAGAA ACAAACAAAG
AGAAAATTAA 43980

AGTTAGATGT AGTGGCACGG TCATAATCTA AAATGTGGCC TAGCTGTTCT
CTGTTCTCTG 44040

TTTCTCTTC TTCCTCCCTC CCTCTCTCTT CTTCATTGTC TGTCTGTCTG
GTGCTTGTAT 44100

ATCAAAATGT AAGTTCTAAG ATATGCTTCA GCACCGTGCC TGCCTGCCTG
CCGCCATGCT 44160

CCACCATGAT AGTCATAGAC CCACCCCTCTC GAACTGTGAA TCCCAAATTT
ACTTTCTTCT 44220

ATGAGTTGCC CTGGTTATGG TGCCTTATCA CAGCAACAGA GCAGTGAGTA
ATATAACCCAC 44280

CCTCAAAGAC AAGCTGAAAG AGAGACCCAT GTGCTGTGGC ATGCGTGTGC
CTACACTTAA 44340

CACACATAAA TAAATACATC TCCTGAAGAA AATTAAAAG TTATTCTGGA
CAGAAACTAG 44400

AGAGGCCAGA CTGGCCTCAG CTCAAGCCCA CAGCAGCTCC TCTGTCCTGC
TGTCCCTTCC 44460

TGTAGAGAAA TTCAGTGAGA CCCAAGCTGT CTGTCCTAGG GCTATAAGCT
GGGTGGGTGG 44520

CTGGGATGAC CACACTTGAT AGAAAAGAGG AAAAGGAACT GGGAGTTGCG
GCCGCC 44576

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GGACAGCCCC AAGGACTACA GGT 23

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CGAAGAACTC CGCAGGGTCC 20

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 18 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AAGACCCGCC ACGACCCG 18

(2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GAATCAGCAC CCTCTCCGCC 20

(2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

TGCGGAGTTC TTCGTGCTGA TGGAG

25

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GGTGCTCGGC GGCGTCCTTC

20

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GAGTGGCGGA GAGGGTGCTG A

21

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GGCCGAGGCT GAGCGGGG

18

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

CTGAAGGACG CCGCCGAGCA

20

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

CTCCAACGCC TGCCGCTGC

19

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GCAGGAGGAG CGGGAGCAGG A

21

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic Primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

TCCAGTGCCC CGCAAGCCG

19